Budget Estimating Models for the Construction Cost of Public School Buildings in Saudi Arabia

by

Tariq Abdul Basset Mohammed Shehab Eldeen

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

CONSTRUCTION ENGINEERING AND MANAGEMENT

July, 1996

BUDGET ESTIMATING MODELS FOR THE CONSTRUCTION COST OF PUBLIC SCHOOL BUILDINGS IN SAUDI ARABIA

BY

TARIQ ABDUL BASSET MOHAMMED SHEHAB ELDEEN

A Thesis Presented to the

FACULTY OF THE COLLEGE OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

CONSTRUCTION ENGINEERING AND MANAGEMENT

JULY, 1996

UMI Number: 1384109

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

COLLEGE OF GRADUATE STUDIES

This thesis, written by **MR. TARIQ ABDUL BASSET SHEHAB ELDEEN** under the direction of his Thesis Advisor and approved by his Thesis Committee, has been presented to and accepted by the Dean of College of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN CONSTRUCTION ENGINEERING & MANAGEMENT.**

THESIS COMMITTEE

Dr. Ali A. Shash (Chairman)

Dr. Sadi Assaf (Member)

Dr. Soliman, Almohawis (Member)

19/11/96

Dr. Al-Khalil, M (Member)

25/11/96

Dr. M. Osama Jannadi
Chairman, Department of Construction Engineering & Management

Dr. Abdullah M. Al-Shehri
Dean, College of Graduate Studies

COLLEGE OF GRADUATE STUDIES

25/12/1996
Date

بسم الله الرحمن الرحيم

Dedicated to my parents

&

to my wife.

ACKNOWLEDGMENTS

Acknowledgments are due to King Fahd University of Petroleum and Minerals and to King Abdulaziz City for Science & Technology (KACST) for providing support for this research.

I wish to express my respect and appreciation to my adviser, Dr. Ali A. Shash, who served as the Thesis Committee Chairman, for his excellent guidance and support throughout this work. I also wish to thank the other members of this thesis committee, Dr. Soliman Almohawis, Dr. Sadi Assaf and Dr. Mohammad Al-Khalil for their constructive comments and suggestions. Special thanks are due to Dr. Walid Al-Sabbah from the Department of Mathematical Science, for his guidance and in-depth discussion on the subject of statistical analysis. Thanks are also due to Mr. Yousef Al-Yousef from the Information Technology Centre at KFUPM for his guidance and help in working with the SAS package.

I express my special gratitude and will remain ever in debt to my parents for their continued support and encouragement. I also express my appreciation to my wife for her understanding and unlimited patience throughout my study.

THESIS ABSTRACT

FULL NAME OF STUDENT : Tariq Abdul Basset Mohammad Shehab Eldeen

TITLE OF STUDY : Budget Estimating Models for the Construction Cost of Public Boys School Buildings in Saudi Arabia.

MAJOR FIELD : Construction Engineering and Management

DATE OF DEGREE : July, 1996

The Saudi Government encourages pre-college education and allocates a generous part of its budget every year for enhancing its quality and for making it available for all citizens. The Ministry of Education implements the governmental policy towards educations and builds hundreds of schools every year. In order to build these schools, the Ministry of Education has to prepare the needed budget and submit it to the Ministry of Finance for its approval. The submitted budget has to be accurate enough to prevent problems associated with underestimation or overestimation of the project cost, such as cancellation of the project at an intermediate stage or building the project below the desired standards. So, in order to help the Ministry of Education to prepare a reliable cost estimate for its school projects and achieve its predetermined plans, this research was conducted. In this research, an identification and evaluation of the budget estimating method currently used by the Ministry of Education were carried out.

This evaluation process revealed major strengths and weaknesses of the system. In order to upgrade such a system, mathematical models were developed. These mathematical models were designed to produce the budget estimate at the conceptual stage when only little information are available about the project. The predictability of the developed models was checked and found to produce better estimates than those produced by the current estimating method used by the Ministry of Education. Moreover, the outputs of the developed mathematical models were checked and found to comply with the accuracy range recommended by the American Association of Cost Engineers (AACE).

MASTER OF SCIENCE DEGREE

KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS

DHAHRAN, SAUDI ARABIA

JULY, 1996

ملخص الرسالة

اســم البــاحــث : طارق عبد الباسط محمد شهاب الدين
عنوان الرســالة : نماذج رياضية لحساب تكلفة إنشاء مشاريع المدارس الحكومية في المملكة العربية السعودية
التــخــصــص : هندسة وإدارة التشيد
تاريخ منح الدرجة : يوليو ١٩٩٦م .

لقد أولت حكومة المملكة العربية السعودية إهتماماً كبيراً بالتعليم وخاصة التعليم ما قبل المرحلة الجامعية حيث وفرت له الإمكانيات الضخمة والميزانية اللازمة التي جعلته ميسوراً لجميع المواطنين ، وإنطلاقاً من تنفيذ سياسة الدولة نحو رفع مستوى التعليم فقد قامت وزارة المعارف بالمملكة العربية السعودية بإقامة المئات من المدارس كل عام . ولكي تتمكن وزارة المعارف من تنفيذ خططها في إنشاء العديد من المدارس سنوياً كان لزاماً عليها تحديد التكلفة اللازمة مسبقاً وتقديمها إلى وزارة المالية لأخذ الموافقة عليها . لذا يجب تقدير هذه التكلفة بقدر عالي من الدقة لتفادي حدوث الإحتمالات التي يمكن مواجهتها نتيجة عدم الدقة في حساب تكلفة المشاريع المزمع إقامتها مثل الغاء المشروع قبل استكماله أو تشييد المشروع بمواصفات ومقاييس أدنى من المطلوب . وبغرض مساعدة وزارة المعارف في تطوير نظام حساب تكلفة مشاريع المدارس التي تقوم بإنشائها فقد تم القيام بهذا البحث .

في هذه الدراسة قام الباحث بإلقاء الضوء على النظام الحالي والمتبع من قبل وزارة المعارف في تقدير تكلفة إنشاء مشاريع المدارس الحكومية في المملكة وتقييمه حيث لوحظ بعض نقاط القوة والضعف في هذا النظام . وبناءً عليه فقد قام الباحث بتطوير النظام الحالي عن طريق استنتاج نماذج رياضية يمكن بها حساب تكلفة تشييد المدارس الحكومية التي لم تنتهي بعد مرحلة تصميمها حيث تعتبر جميع المعلومات المتواجدة معلومات مبدئية وليست تصميمات نهائية . تمت المقارنة بين النتائج التي تم الحصول عليها باستخدام النماذج الرياضية المستنتجة بالنتائج التي قد تم الحصول عليها باستخدام نظام حساب تكلفة المشاريع والمستخدم حالياً من قبل وزارة المعارف حيث تبين تفوق النماذج الرياضية المستنتجة على النظام الحالي لوزارة المعارف من حيث الدقة . كما تمت مقارنة نتائج هذه النماذج الرياضية بمستوى الدقة المنصوص عليه من قبل الجمعية الأمريكية لمهندسي الكلفة حيث تبين أنها في حدود المسموح به.

درجة الماجستير في العلوم
جامعة الملك فهد للبترول والمعادن
الظهران ، المملكة العربية السعودية
يوليو ١٩٩٦م

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

The preparation of a literate population requires a sincere devotion to the development and maintenance of the three major components of pre-college education: curriculums, teachers, and physical facilities.

In Saudi Arabia pre-college education is available for the people in villages, cities, and remote areas. The Saudi Government works hard to increase the literacy level of its nation. Figures 1,2 and 3 show the number of students enrolled in elementary, intermediate and high school levels, respectively, in 1990.

This pre-college education is provided for all citizens by the Saudi Government and a very few private investors. Figures 4,5,6 show the number of elementary, intermediate and high schools, respectively, built until 1990.

The Government in Saudi Arabia allocates a generous part of its national budget for pre-college education (Table 1) in an effort to improve the nation's literacy level. Since the Ministry of Education is one of the governmental agencies which takes care of enhancing the education process in the Kingdom of Saudi Arabia through increasing the number of schools or encouraging more students to be enrolled in schools, to improve its efficiency in any of its aspects such as cost estimating will be of great importance and help to the Ministry itself as well as the Saudi government.

In this research a review will be made of the estimating method that is currently used for the construction cost of school building projects by the Ministry of Education and mathematical models will also be developed to predict the construction cost of school projects in the Kingdom of Saudi Arabia.

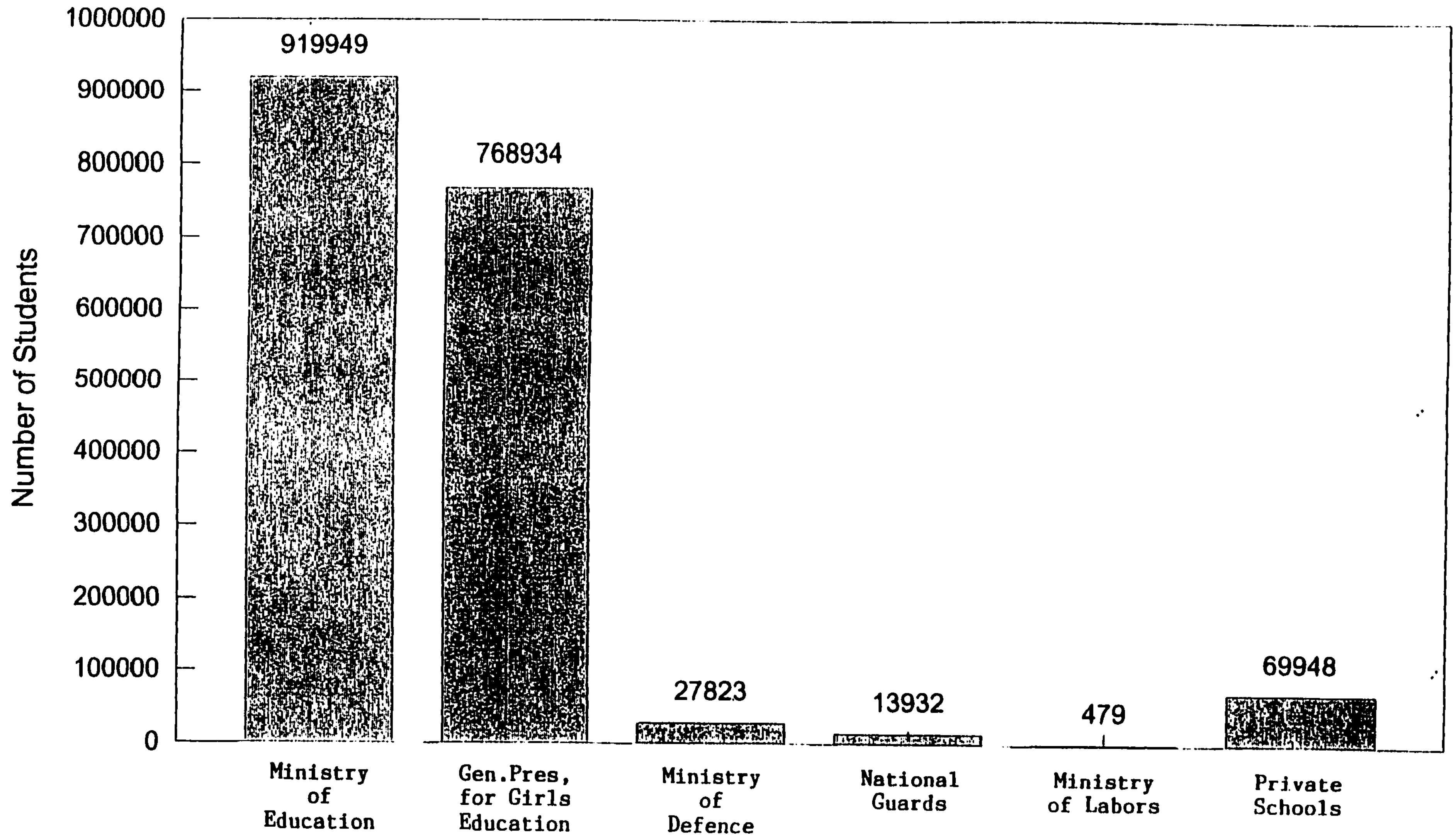

Figure-1 Total Number of Students at Elementary Level (1990)

Annual Statistics Book, 1991

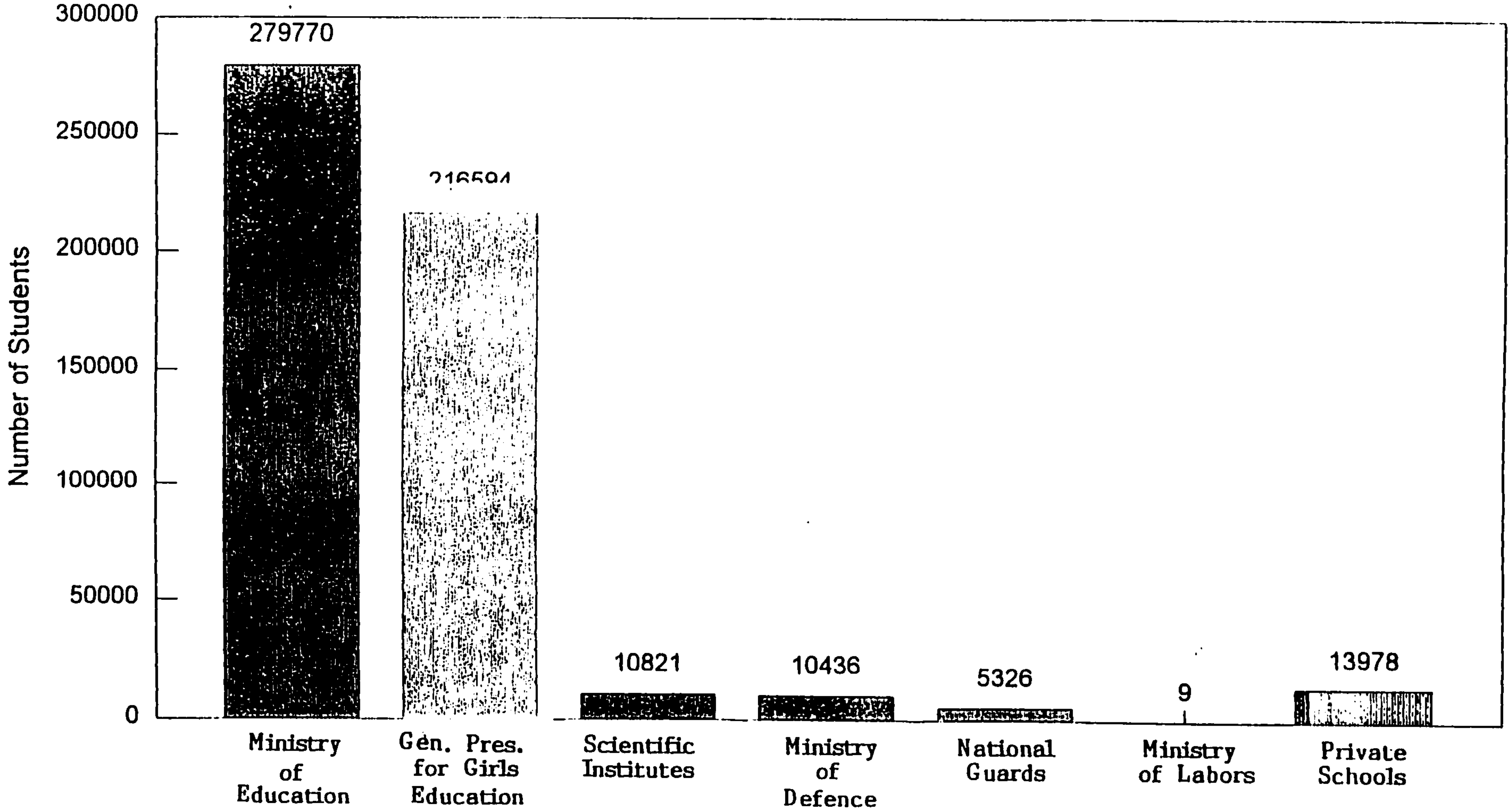

Figure - 2 Total Number of Students at Intermediate Level (1990)

Annual Statistics Book, 1991

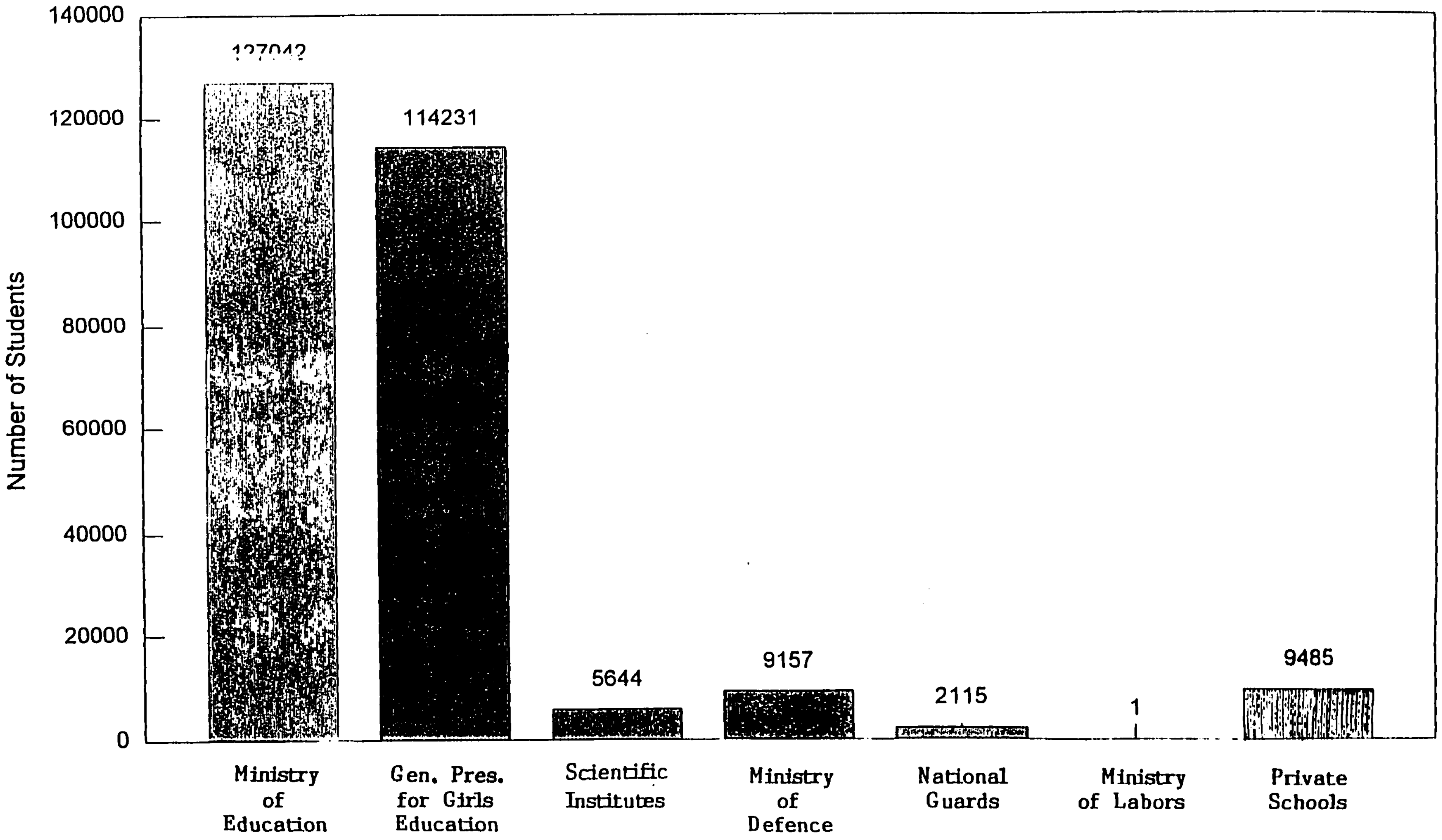

Figure-3 Total Number of Students at secondary Level (1990)

Annual Statistics Book, 1991

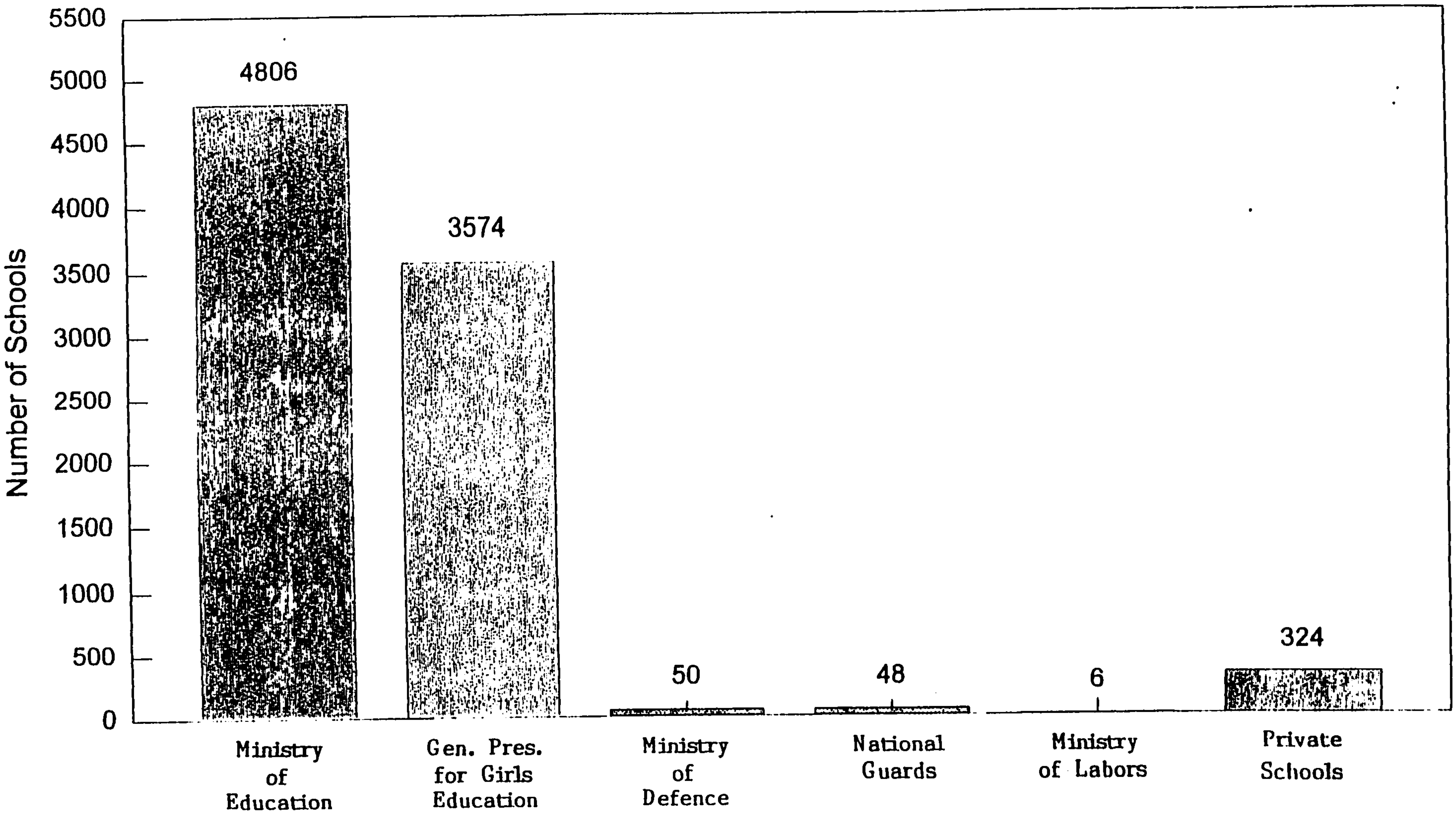

Figure-4 Total Number of Schools at Elementary Level (1990)

Annual Statistics Book, 1991

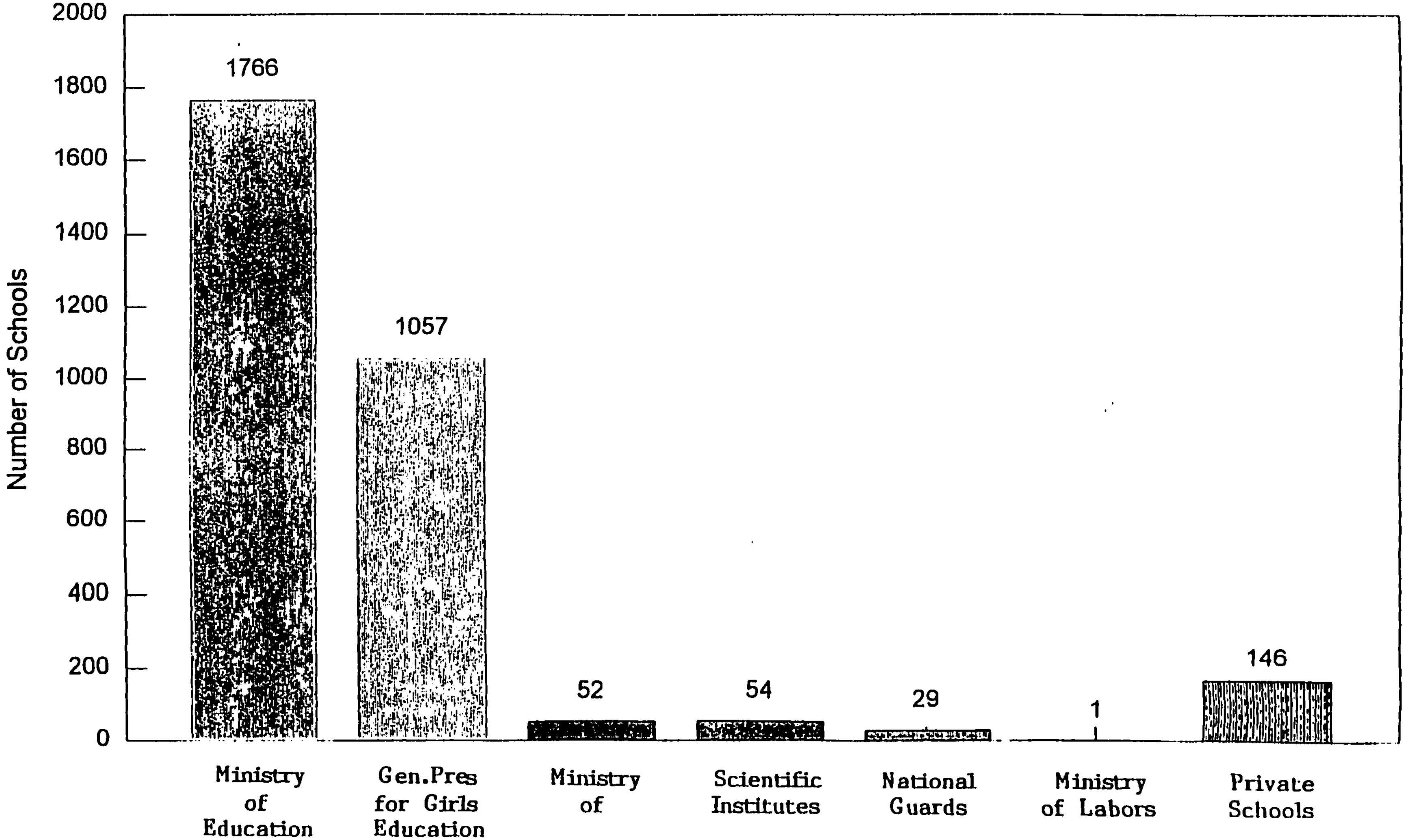

Figure-5 Total Number of Schools at Intermediate Level (1990)

Annual Statistics Book, 1991

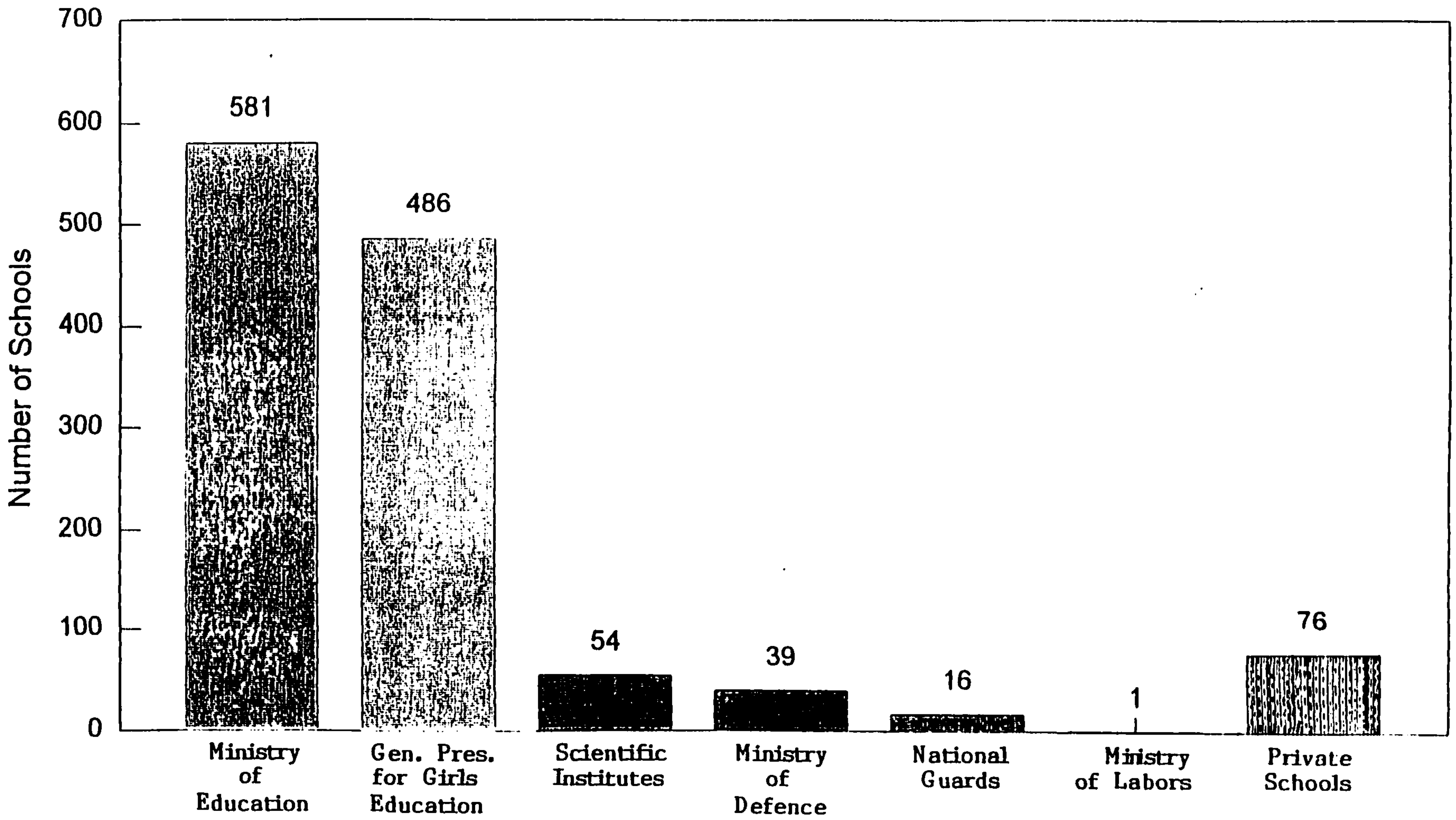

Figure-6 Total Number of Schools at Secondary Level (1990)

Annual Statistics Book, 1991

Table -1

ANNUAL GROWTH OF GOVERNMENT APPROPRIATION FOR EDUCATION
1406/1407 - 1410/1411 A.H.
(Million Riyals)

Agency	1410/1411	1409/1410	1408/1409	1407/1408	1406/1407
Ministry of Higher Education	401.8	454.0	454.6	502.7	426.6
King Abdulaziz's Darat	8.6	7.9	7.7	8.3	7.1
Ministry of Education	10139.8	9515.0	9064.5	8770.4	7346.0
King Saud University (Riyadh)	1430.0	1357.9	1350.5	1365.2	1283.0
King Fahd University of Pet.	329.3	315.6	317.3	301.5	263.8
Girls Education	7962.5	6842.7	6470.3	6116.4	4828.0
Capital Model Institute	12.9	12.7	12.3	12.5	11.8
Girls Training College	467.1	327.5	311.9	317.6	275.0
King Abdulaziz University	869.9	748.0	750.2	730.1	666.5
King Faisal University	348.4	357.7	354.6	363.7	314.2
Mohammad Bin Saud Islamic University	688.0	617.1	596.0	624.2	524.0
Oum-Al-Qura University	387.3	351.4	352.3	368.3	348.3
Islamic University	180.5	173.4	173.9	177.5	163.0
Typical Model Schools at Jeddah	15.3	15.5	15.3	15.2	14.7
Total	23241.4	21096.5	20231.4	19674.6	16472.0
Education Projects	1342.9	1408.4	2008.4	2754.3	3153.3
Grand Total	24584.3	22504.9	22239.8	22428.9	19625.3

Reference: Ministry of Finance and National Economy

1.1 STATEMENT OF THE PROBLEM

The population in the Kingdom of Saudi Arabia is in exponential increase and accordingly so is the number of students. The government builds hundreds of elementary, intermediate and secondary schools and spends tens of millions of Saudi Riyals every year to accommodate the increasing number of students.

Figure 7 shows the number of students in the three levels of general boys' education. It can be seen that the number of students in the elementary level is almost three times the number of students in the intermediate level, and the number of students in the intermediate level is almost twice the number of students in the secondary level. Consequently, the educational authorities such as the Ministry of Education are expected to build more and more intermediate and high schools to accommodate the students who will be moving from the elementary schools to higher levels of education.

One of the important steps that the school building process passes through and that has a great influence on the project through out its execution time is the preparation of a reliable budget estimate. This is due to the fact that preparing a reliable budget estimate usually saves a lot of problems that might be encountered during the execution process of the project.

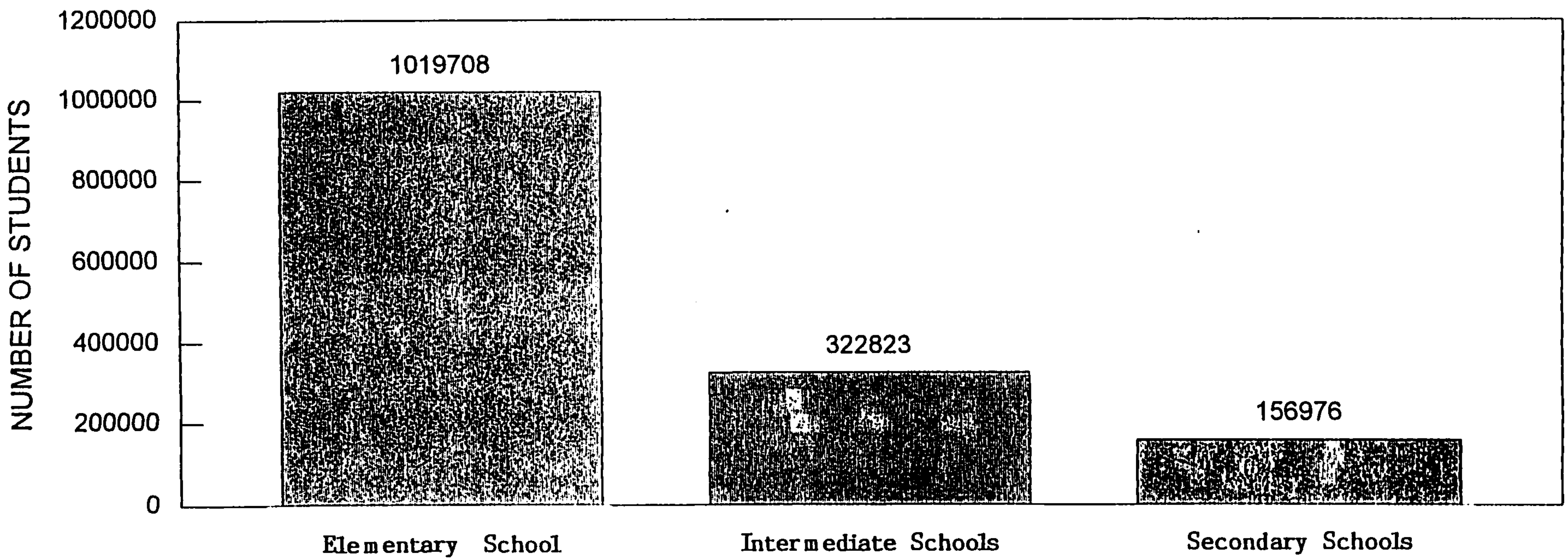

Figure -7 Total Number of Students at Elementary Level (1990)

Annual Statistics Book, 1991

The production of a reliable budget estimate for the construction of schools in Saudi Arabia is important to the Ministry of Education because underestimating the needed budget may cause one or more of the following problems:

1. Delay in awarding school projects until more funds are approved by the Ministry of Finance.
2. The scope of school projects is reduced to satisfy the allocated budget.
3. Building schools below the desired standard.
4. Cancellation of the project.
5. Fewer schools are built.

On the other hand, overestimating the needed budget may cause inefficient utilization of Ministry of Finance's resources.

Therefore, developing a good understanding of the cost estimating system that is currently being used by the Ministry of Education and improving it if possible will be of great importance to the Ministry of Education and the Ministry of Finance. This is because a reliable cost estimate system will help the Ministry of Education to achieve its predetermined plan and the Ministry of Finance to save a considerable amount of money that can be used by the government for fulfilling other needs.

This study is an attempt to answer the following research questions:

1. How does the Ministry of Education estimate the budgetary construction cost for its school buildings?

2. What is the Ministry of Education's opinion on the current system used for estimating the budgetary cost for the construction of its school building projects?

3. How can the current system be improved?

1.2 OBJECTIVE OF THE STUDY

The objectives of this study are:

1. Identify and assess the method currently being used by the Ministry of Education to prepare budget estimates for the construction of its school projects.

2. Develop cost estimate models for the Ministry of Education for estimating the budgetary cost for constructing its school buildings.

1.3 SCOPE AND LIMITATION

By referring to figures 1 to 6, it will be noticed that both the Ministry of Education and the General Residency of Girls Education are the two largest organizations building and operating schools in the Kingdom of Saudi Arabia. It will also be noticed that these two organizations are by far the largest in taking care of boys' and girls' Education in Saudi Arabia.

Due to cost and the time constraint of this research, it becomes necessary to limit the study to one of the two organizations. The Ministry of Education is selected since it is the larger of the two and because initial contracts have already been established.

Schools built by other agencies such as Saudi ARAMCO and the Royal Commission will be disregarded in this research because they have different standards of construction and accordingly they might introduce some disturbancy in the obtained results.

1.4 SIGNIFICANCE OF THE STUDY

By knowing and appreciating the amount of problems that might face the Ministry of Education due to over-and under-estimation of its projects' costs, the need to study the current estimating system used by the Ministry and to improve it if possible becomes more and more essential. This is to avoid the problems that might be encountered with inaccurate estimating of the project cost and to get the maximum utilization of the allocated budget for building new schools.

Over-and under-estimations of the project cost can be considered as one of the causes of exhaustion of resources and stoppage of the project at an intermediate stage respectively. If a reliable budget estimate system is available, then the problems associated with over-estimation and under-estimation of the project cost will be reduced and the potential for achieving the following targets will be maximized:

1. Minimizing the delay associated with securing additional appropriation from the Ministry of Finance.
2. The school project will be constructed as per the predetermined plans and specifications . No reduction in the scope of work will be made during the progress of the project.
3. No stoppage of the project at an intermediate stage.

4. Only reliable contractors will be involved in the bid, ensuring strong competition.

5. Meeting the already planned number of schools to be built within a specific period.

6. Maximum utilization of financial resources.

7. Creating a common base for estimating the project's budget by the Ministry of Education and the Ministry of Finance.

CHAPTER TWO

LITERATURE REVIEW

A construction cost estimate can be defined as the best judgment of what a project will eventually cost (Keith Collier, 1987). This judgment of the project approximate total cost is usually based on information obtained from a specific set of drawings and its related specification, past experience, completion of previous similar projects, etc. As the estimate is made before the work is done, the agreement of the estimated cost with the actual cost will depend upon the skill and judgment of the estimator. Skill implies the accurate use of proper estimating methods, and judgment implies the correct visualization of the work as it will be done (Raddon, P. S.B., 1982).

Estimates for construction costs are used for different reasons and so are made by different methods and provide different answers. Some methods are adequate for the purpose intended and are relatively simple. Others are more detailed and require more time. Deciding upon which cost estimate method is to be used is dependent on the purpose of the estimate, the time available for making the estimate, the amount of information required by the system and the degree of accuracy which the estimator is looking for.

Four different parties are usually involved in the development process of any construction project. These parties are an owner, a designer, a contractor and a financing authority (if any). The involvement of each party starts at a certain stage to serve a specific need.

The owner estimate is different than those estimates made by the contractor and the designer because it is different in form, function and technique. In case the owner is a private entity or he is constructing a commercial project, he makes his estimate mostly for a feasibility study and for funding purposes. On the other hand, if the owner is a public entity or he is constructing a project for serving the community without expecting any financial returns, he prepares his estimate mainly for budgeting and funding purposes. Of course, other purposes such as selection of alternative design proposals and control of cost during the design phase are objectives sought by both classes of owners.

The techniques (methods) of cost estimating vary greatly. Focusing on budget estimating, conceptual or preliminary estimates are usually performed. The literature contains many studies on methods for preparing conceptual or preliminary estimates for construction projects. Among them are those included in the following paragraphs.

2.1 PRELIMINARY ESTIMATING

2.1.1 The Unit Cost Estimate

Unit cost estimate are the quickest and easiest preconstruction estimates to prepare. This type of estimate can be broken down into 5 different types of estimates depending on the base elements considered. The base element can be one of the following:

- Function
- Floor Area Units
- Building Volume Units
- Enclosed Area
- Trade

2.1.1.1 The Function Estimate

The function estimate measures the cost of a building relative to its use or function. This can be done by determining the historical cost per student for a school building or the historical cost per bed for a hospital or the historical cost per seat for a theater, etc. Then, multiplying these historical costs by the quantity of the function element corresponding to each project yields the total cost of the project.

2.1.1.2 *Floor Area Unit Estimate*

This is based on the historical cost per unit floor area data collected by the estimator. By multiplying the historical square foot cost by the calculated square footage of the floor area for the proposed building, a preliminary cost estimate can be determined.

2.1.1.3 *Building Volume Unit Estimate*

This type of estimate relates the cost of a building to its volume. The historical unit cost per cubic foot of a building is collected. Multiplying this historical unit cost by the volume of the proposed building yields the total cost estimate.

2.1.1.4 *Cost Per Enclosed Area Estimate*

This type of estimate is based on the area of all the horizontal and vertical planes of the building. In this type of estimate the area of the floors is added to the interior areas of the walls. The historical cost per the sum is collected. Multiplying the historical unit cost by the total area of floors and walls of the new proposed building yields the total cost estimate of the project.

2.1.1.5 Trade Unit Cost Estimate

This type of estimate breaks down the total building into the basic parts or trades. Past unit cost data of these trades are used to make the cost estimates.

In order to estimate the cost of a project using this technique, the estimator collects the past costs for different trades such as excavation, concrete work, masonry work, doors and windows, and so on. The estimator then relates the past costs of those trades to their unit volume or area. These determined past costs per unit volume or area are then used to estimate the cost of the new project by simply multiplying them by the unit volume or area of each trade for the new project.

2.1.2 Parameter Estimate

This method depends on developing historical cost summaries for various building's components such as concrete, masonry, plumbing , etc. and relating these costs to the physical parameters of the same building such as the gross enclosed area, net finished area, roof area, etc. The relation is simply found by dividing the total cost of the components (parameters) by the physical parameter (area or volume) in order to obtain the cost per unit of that parameters. These unit costs are then multiplied by the physical parameter measurements of the proposed new building to obtain the total cost estimate. Care must be taken in choosing the parameters of the old and new buildings in the sense that they must be the same. Also the major cost areas for both buildings must be the same; in other words, this method can be used only for similar projects.

2.1.3 Factor Estimate

This type of estimate is best used for projects with a predominant cost component. Usually this component is the equipment purchased for the project. Examples of such projects are oil refinery, petrochemical plant, power stations, etc. In this type of estimate, the total project is broken up into a series of components such as equipment costs, equipment installation cost, process piping, and so on. After listing these components for previous projects the cost of the major equipment is given a factor of one. The factors for other items are developed by dividing their corresponding costs by the cost of the major equipment. By collecting data for numerous similar previous projects, the reliability of the obtained factors becomes greater. For estimating the cost of the new project, the updated cost of the major equipment is only determined. These historical obtained factors are then multiplied by the major equipment cost to calculate the cost of other components. Adding all those costs together yields the total cost of the project.

2.1.4 Range Estimate

All previously mentioned methods of estimating utilize a single point approach in order to determine the future cost of a particular project. No matter how much experience goes into developing this single point estimate, it is highly unlikely that the actual value will fall precisely at that number. Range estimating has the objective of setting out various project costs within this range. In other words, range estimating calculates how much higher or lower the actual cost varies from the single point estimate.

The range estimating process does not limit itself to an estimate of a single cost for a project. Instead, the use of the process states a target cost, a lowest estimated cost, a highest estimated cost and a confidence limit that the actual cost will be equal to or less than the target cost. These different types of project total cost are calculated based on the project's individual work phases or packages. The result is that the user of the range estimating process can equate risks with various possible project budgets. For example, a range estimate may indicate that there is an 80% probability that a project will cost less than 1 million dollars, an 85% probability that it will cost less than 1.1 million, and so on.

2.2 REGRESSION ANALYSIS

Regression analysis is not really a cost estimating technique but it is considered as a mathematical tool for performing data analysis. It is considered as a good technique for developing statistical models and it has proved its strength in cost estimate discipline since a long time ago. For this reason it will be briefly discussed in the following paragraphs.

Regression is described as the study of relationships between quantitative variables. It is often used to predict the response variable from the knowledge of the independent variables. Among the variables studied, the variable of major interest (in which the change is to be studied) is designated as the response or dependent variable, and the variables that have significant effect on the response variable are termed as independent variables.

There are generally seven steps to be followed in constructing a regression model. The steps followed in the model - building and estimation process are given below (Bozai, G. A, 1981).

Step 1: Identify the variables to be included in the regression model.

As a first step towards building the regression model, the dependent variable and the independent variable (s) that will be used in constructing an estimating equation are identified. A great deal of time is spent in identifying independent variables, or variables that can be used to estimate values of the dependent variable. It is generally not known until step 4 is completed whether it will be necessary to include other independent variables in the estimating equation.

Step 2: Collect sample data.

This step consists of collecting all needed information on all independent variables that were identified in step 1. The data sources must be solid and reliable, such as written documents, or even verbally acquired through interviews.

Step 3: Specify the relationships that exist between the dependent and the independent variables, and between the independent variables.

The nature of the relationship between the dependent and the independent variables has to be specified. A plot of data points collected in step 2 provides clues to the form of the relationships between the variables of interest.

Step 4: Determine model parameters.

This step consists of the determination of the parameters of population regression model using the sample data.

Step 5: Verify assumptions.

The determination of whether the underlying assumption of the model has been met is a complex task. The construction of residual plots is often quite helpful in an initial determination of whether the assumptions have been met, or whether it will be necessary to identify new variables to be included in the model or to specify a different functional for the manner in which the variables interact. This may lead the researcher back to step 1 and through the process once more if it is discovered that the model developed is not appropriate or that the underlying assumptions have been violated.

Step 6: Statistically test the usefulness of the model developed.

At a stated or desired level of significance, the model developed or the terms included in the model are tested for statistical significance. This is a part of the hypothesis testing process.

Step 7: Use the model developed for prediction and estimation.

Once a regression model is developed, it is desirable to use the model for estimating or predicting the values of the dependent variable.

Regression analysis has proved to be an indispensable tool for data analysis in a variety of disciplines. One of those disciplines is cost estimating. Regression analysis has been used extensively in preparing cost estimates for different projects, such as highway construction and water supply projects. In the case of highway projects, a set of multiple linear regression models were developed for different categories of highway construction and maintenance projects for the state of Alabama highway department. Practical applications of the developed models resulted in a range of accuracy of 2 to 33% of the actual cost. In the case of water supply projects, a regression analysis model was developed to estimate the cost of water supply projects for the Sultanate of Oman. The developed model was able to predict the cost within an average error of 0% to 14% (Bozai, 1981 & Mashoor, 1993).

CHAPTER THREE

RESEARCH METHODOLOGY

This chapter presents all the necessary steps that were followed to achieve the two objectives set for this study. Before going into the details of this research methodology, it will be very helpful to go over Figure 8 in order to give a brief introduction to the steps that are necessary for achieving the research objectives.

As stated previously there are two objectives for this study; the first is to assess and evaluate the current estimating system used by the Ministry of Education, and the second is to develop statistical models to calculate the budget estimate for building school projects for the Ministry of Education. To achieve the first objective, the data needed can be grouped into four main categories, which are method description, characteristics of the system, advantages and finally disadvantages. This data will be collected through interviewing the Head of Project Studying Departments. Based on the collected data, the system will be evaluated. This evaluation will be either positive or negative and in both cases it will lead to the development of a new budget estimating models for the Ministry of Education.

For the development of the new estimating models, the parameters that are significantly affect the construction cost of school projects have to be determined. These parameters were determined from the literature, the available building designs for the Ministry and the output of the evaluation process of the current estimating system. The data and information on those parameters were collected from the archives of the Ministry for a predetermined sample of schools. These collected data were statistically analyzed to develop the needed models. The developed models were verified by selecting some schools other than those used for developing the models to compare their bid price to the budget estimate determined by the models.

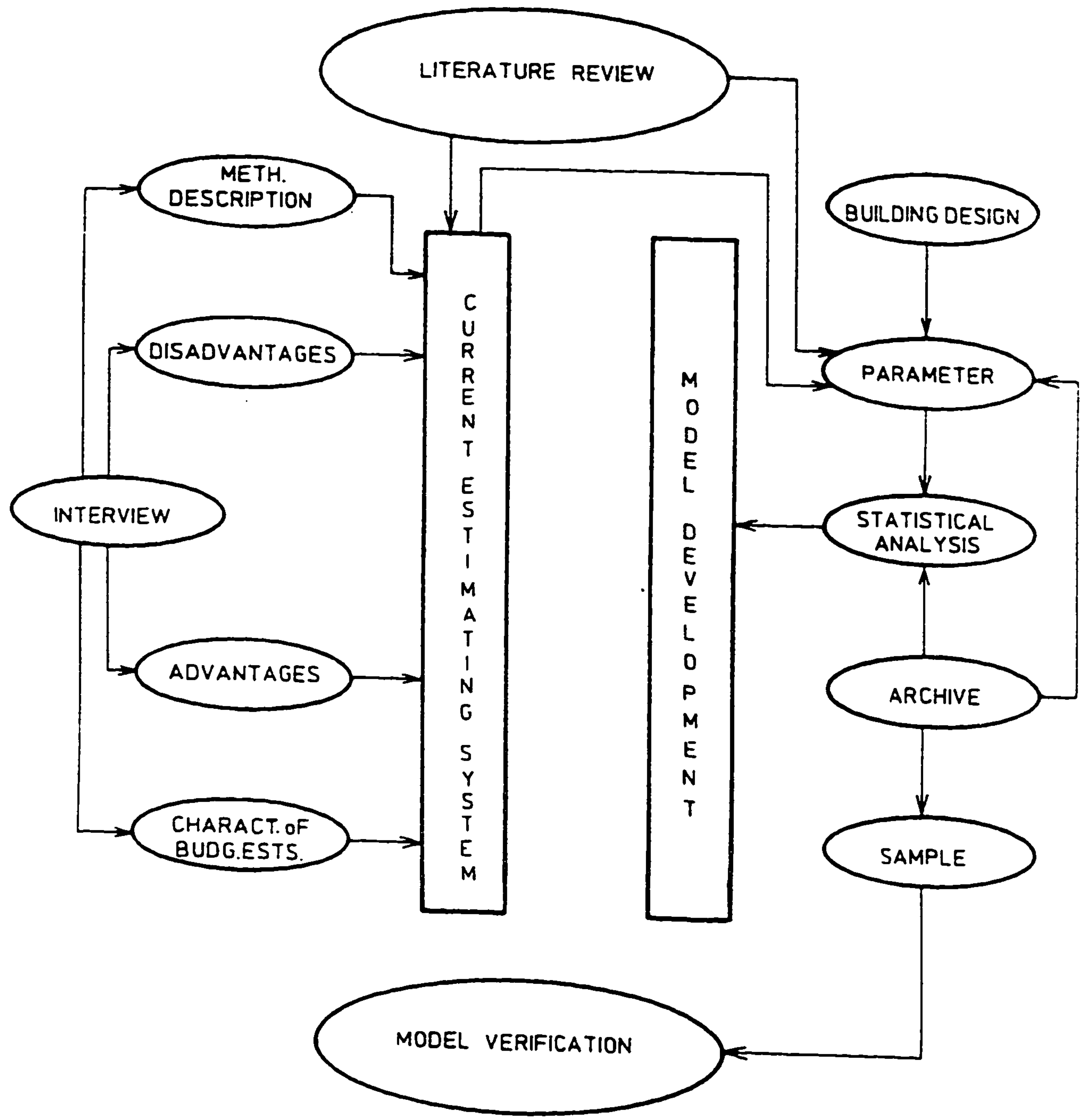

FIG. 8 : RESEARCH METHODOLOGY

3.1 CURRENT ESTIMATING SYSTEM

3.1.1 Required Data

The following information was required for investigating the current budget estimating system used by the Ministry of Education:

1. How the budget estimate of school projects is prepared.
2. The type and characteristics of the Budget Estimate System that is currently used by the Ministry of Education:
   - when was it developed ?
   - who developed it ?
   - what are its input data and basic components ?
   - is the system computerized or manual ?
   - who performs the estimate ?
   - is it one system for all types of schools or different for different levels ?
   - who approves the estimate and what are the criteria for approving ?

- if the system is manual, is there any standard format to follow ?
- does the system give a lump sum price or separate prices for major items such as concrete work, excavation, etc. ?

3. Advantages
4. Disadvantages

3.1.2 Data Collection

The required data were collected by visiting the Project Studying Department in the Ministry of Education. An interview was conducted with the head of the Department. The purpose of this interview was to collect as much information as possible about the current estimating system used by the Ministry of Education. Basically this interview concentrated on achieving an overview of the complete estimating system that is currently used, such as its history, process, strength, weakness, its components and any other information needed to enhance the evaluation process of the system. During this interview an open-ended questionnaire, as shown in Tables 2 to 5 in Appendix A, was used to facilitate the process of collecting the needed information and also to make the interview as efficient as possible.

3.1.3 Data Analysis

The first objective of the study was achieved by discussing the existing system in detail to describe the following:

1. How the current Budget Estimating System works.
2. The system in general:
   - what are the steps followed in preparing the budget estimate?
   - what are the steps followed to get approval for the budget from the Ministry of Finance?
3. How it was developed.
4. Its input data and main components.
5. Advantages and disadvantages.
6. The nature of the system and the end results.

Based on an evaluation of the information obtained from the interview and the resulting potential areas of improvement, a proposal was made to develop new statistical models that are able to avoid most of the problems encountered by the existing budget estimating system and to provide a cost estimate that is as close as possible to the bid price submitted by the contractor.

3.2 DEVELOPMENT OF THE STATISTICAL ESTIMATING MODELS

3.2.1 General Approach

The study will develop a statistical model to assess the construction cost of schools (CC) in which our knowledge about the school environment (V) and physical characteristics (P) prior to the commencement of the design process will be used. V includes our knowledge of the environment in which the schools are constructed, i.e. location, type of soil, etc, P includes our knowledge of the physical characteristics of the school such as size, number of class rooms, etc.

Suppose that construction cost, CC, in a particular school project is affected by different physical and environmental factors such as size, location, net enclosed area, etc. These factors are measured by their level X. The level of factor K is X_k. For example, if the K factor that affects the construction cost (CC) is the number of students and the number of students is 500, then $X_k = 500$ students. The level of various physical and environmental characteristics, X_k, would have different effects on CC. Then each level will be multiplied by an appropriate weight, β_k, and the total various weighted factors will be summed to obtain the total effect on CC.

$$CC = \beta_o + \sum_{K=1}^{n} \beta_k X_k + E$$

where

CC = Construction cost of school project

β_k = regression coefficient (parameter) for independent variable X

X_k = The level of factor K

E = Random scatter component error term

β_o = The Y - intercept on Y axix.

The regression expression is read as the construction cost , CC, is equal to the sum of all levels of the school's physical and environmental characteristics, X_k, times the weights, β_k, plus an error term E which has a probability distribution center at zero and a finite variance. The assumptions are made that the variance is constant over all school construction costs and the factor weights, β_k's, remain fixed on all schools.

The development of such a model for CC requires, firstly, the determination of the physical and environmental factors that significantly correlate with CC. Secondly, assuming that the past environmental situation will persist and be repeated in the future, the use of information from past school projects can be employed to predict future construction cost. It is likely that the future of schools construction will be a continuation of the current and past situation. The past information allows the calculation of the factors (Predictors) weights, βk, and the variance, σ^2. The weights of parameters in the above equation were found by using the least square method (Appendix D).

3.2.2 Required Data

The required data for the developing the statistical models for calculating the construction cost of boys' schools projects can be grouped into two categories. First, the dependent variable and second is the independent variables. In the following sections, both categories will be explained and discussed.

3.2.2.1 Dependent Variable

As discussed earlier, the dependent variable in these statistical models that will be developed for calculating the construction cost of school projects is the cost of the project. Since the objective of this study is to predict the budgetary cost for school buildings built by the Ministry of Education which is considered in this case as an owner estimate, the bid price was considered. This bid price is the price submitted by the contractor for constructing the concrete sckelton of the building, finishing activities, foundation work lighting system, plumbing system and all site preparation activities.

3.2.2.2 Independent Variables

The cost of any construction project is a function of several factors. In order to come up with the most influential factors in determining the construction cost of a school project, all design models that are available with the Ministry of Education were analyzed and reviewed with the objective of determining those influencing factors. In addition, a close consultation was carried out with the Project Studying Department and the Geotechnical Department in the Ministry of Education to determine the most common cost variation sources for school projects. The feed-back obtained from the assessment process conducted on the current estimating system was also utilized to come up with some factors that are heavily affecting the construction cost of school projects.

A literature review was also conducted on many cost estimating books and manuals. Few of them solely discussed cost estimating of specific projects such as school buildings, while many of them talked about cost estimating for buildings in general. Literature divided the potential cost factors in any building as follows:

- Site work and execution.
- Concrete Work.
- Masonry.
- Metals.
- Thermal and Moisture Protection.
- Doors and Windows.
- Finishing.
- Mechanical.
- Electrical.

Literature review revealed also that construction year, project's execution time and location of the project might have a dramatic effect on the cost of the project. Other literatures (Dodge, 1981) stated classrooms and technical laboratories as variation sources between different schools depending on their level and usage.

Moreover, 'The Building Construction Cost Data, 1981' used the number of students as a basis for calculating the total construction cost of school projects. This cost also varied depending on the level of school (i.e. elementary, intermediate or secondary).

From the above discussion, it might be noticed that the factors which are thought to affect the construction cost of boys schools were gathered from three main sources. These sources are the feed-back obtained from the evaluation process conducted on the current cost estimation system used by the Ministry of Education, the literature review and finally, studying of the available designs with the Ministry of Education. Each source contributed to the main list of factors that are thought to affect the construction cost of school projects by suggesting one or more factors.

In the case of the feedback obtained from the evaluation of the current estimating system used by the Ministry of Education, it helped in giving more importance to all factors that are related to the ground conditions and the type of footing and accordingly to consider them as governing factors for calculating the cost needed for building school projects. Hence, the foundation type, soil bearing capacity, soil improvement, amount of excavating and backfilling materials were hypothesized to be responsible for the variation in the construction cost of boys schools. The literature review revealed also that the number of classrooms, number of technical laboratories, level of school, construction year, project's execution time, location of project and number of students are governing factors in calculating the construction cost of school projects. Since it was found that the number of students and number of class rooms in those schools built by the Ministry of Education are very highly associated as each classroom is designed to accommodate 26 students, it was decided to consider only the number of class rooms as they are more clearly identified by the drawings and designs prepared by the Ministry of Education. Therefore, the number of class rooms, level of school, construction year, project's execution time, location of the project and number of technical laboratories were selected as governing factors in determining the construction cost of school projects.

Upon studying the available designs with the Ministry of Education, it was found that they are different with respect to their length of fence, concrete area and length of retaining walls and so they were selected as governing factors in calculating the construction cost of school projects.

From the above discussion and bearing in mind that all necessary factors have to be available by the time the budget estimate is made and must also be obtained within a minimum time, the following group of factors are hypothesized as the factors responsible for variation in the construction cost of boys' schools.

1. Enclosed concrete area of school (all areas enclosed by concrete structures, i.e. classrooms, administration offices, labs., guard house, etc.

2. Number of classrooms.

3. Number of technical laboratories.

4. Construction year (contract awarding date).

5. Project's execution time (time taken to build the school, i.e. from the date of receiving the notice to proceed until the school is completely finished).

6. Length of fence.

7. Foundation type.

8. Soil bearing capacity.

9. Amount of excavation material.

10. Location of school (i.e., remote area, in town etc).

11. Level of school: whether elementary, intermediate or secondary.

12. Amount of backfilling material.

13. Length of retaining walls.

14 Soil improvement.

Other variables were thought to have an impact on the total construction cost of the school project, such as the province, area of opened and closed sports courts and type of shading material for the open court, but unfortunately no information was available to the researcher about them.

3.2.3 Identification of Cost Determinant Factors

3.2.3.1 Concrete Area of School

The enclosed concrete areas of a school are basically all areas enclosed and shaded by concrete such as classrooms, administrative offices, laboratories, guard house, etc. As the enclosed area increases, the quantity of building materials, equipment and the number of man hours necessary to construct the project also increase and so does the total cost of the school project. The magnitude of this factor is expressed in m^2.

3.2.3.2 Technical Laboratories and Level of Schools

A technical laboratory is a space designated for experiment and exhibition of scientific items for students such as physics and chemistry and such technical laboratories are not usually available in all types of school. However, if they are part of the school, they will increase the cost of the project. This increase in the cost is due to the additional equipment cost involved in the machines and instruments used.. Another cost area involved in constructing these laboratories is the different requirements of building codes, standards and the special requirements of the Ministry of Education which assures the safety of all students. As the number of these laboratories increases, the cost of the project also increases and vice versa. Usually the number of available laboratories in any school built by the Ministry of Education is dependent on the level of the school itself as defined below.

For elementary school, no laboratories are available

For intermediate school, two laboratories are available

For secondary school, four laboratories are available

The unit of this factor is expressed in absolute numbers (i.e. 0,2 or 4).

3.2.3.3 Construction Year

The cost of any project and manpower also vary depending on the year in which they were purchased or utilized. This variation in cost is mainly due to inflation and other economic factors such the availability of items and the overall local and international market conditions. This factor will be ignored in our analysis since all the bid prices collected in our sample were submitted by different Saudi contractors in 1994 and can be considered already normalized.

3.2.3.4 *Project's Execution Time*

One of the factors which has an impact on the total cost of a project is the project duration. As the project duration increases the total cost decreases up to a certain point where it starts to increase again. This variation in the total cost of project throughout its execution time is due to the different behavior of direct and indirect costs. The indirect cost increases with time while the direct cost decreases with time upto a certain point then it increases again. So, it can be concluded that there is an optimum duration for any project at which the total cost of the project will be minimum. This cost will definitely increase if the project duration will be shortened or expanded. This factor will be expressed in months.

3.2.3.5 *Location of the School Project*

The school project being located in a remote area or in town plays an important role in determining the cost of the project. If the project is located in a remote area, the transportation, mobilization and demobilization costs will be higher and accordingly the cost of the project. Another reason behind increasing the project cost if it is located in a remote area is that the contractor cannot easily exchange or shift labor or machines from one site to another when they are idle.

This factor will be ignored in our analysis since all data were collected on those schools built in town.

3.2.3.6 Type of Foundations

The type of foundation varies depending on two important factors:

1. Load coming from building.
2. Bearing capacity of soil.

As those two factors are interrelated, they imply the usage of one certain type of foundation system rather than another. The most common types of foundation are isolated footings, mat foundation and strip foundation. Each one of these types has different construction requirements and consumes a different amount of construction materials and man hours, which of course affects the cost of the project. To enhance the analysis process, the magnitude of this factor will be expressed by using certain code numbers (i.e. 1 for isolated footing, 2 for mat foundation and 3 for strip foundations).

3.2.3.7 Bearing Capacity of Soil

The bearing capacity has an impact on the total cost of the project since it determines the type of footing that will be used and whether any of the soil improvement techniques will be utilized or not. As the bearing capacity increases, the footing area decreases and accordingly the cost of the project decreases too. The factor will be expressed in numbers such as 1 kg/cm^2, 1.5 kg/cm^2, 2 kg/cm^2, etc.

3.2.3.8 Soil Improvement

Different soil improvement techniques might be used during the construction stage of the project to improve or upgrade the quality of soil in terms of its bearing capacity or its main characteristics in general. The usage of any of the known methods of soil improvement will increase the total cost of the project. During the analysis process, this factor will be expressed using certain code numbers (i.e. 0 of non-improved soil and 1 for improved soil).

3.2.3.9 The Availability of Non-Conventional Structures

Sometimes certain types of non-conventional structures, such as a retaining wall, are needed due to the presence of some problems in the site conditions of the school project. If any of these non-conventional structures are required, the total cost of the project will be increased. The only type of non-conventional structure that is usually needed in the school projects built by the Ministry of Education is retaining walls. The height of these retaining walls ranges between 1.5m and 7m. The magnitude of this factor will be expressed explicitly by stating the height of each wall and the number of linear meters for each height.

3.2.3.10 Length of Fence

One important factor that has an impact on the total cost of the school project is the length and height of fence. As the height or the length of the fence increases, the cost of the project also increases. It has been found that the height of the fence is the same (3m) for all school projects built by the Ministry of Education and accordingly its importance in the analysis will be negligible. On the other hand, the length of fence might be different between two identical design models depending on the area allocated for the project. This factor will be expressed in linear meters.

3.2.3.11 The Availability of Excavation or Backfilling

Normally any project involves some excavation or backfilling activities. These activities are important in order to place the substructure at a certain depth or to keep the superstructure at a certain level depending on the different design requirements. Some- times excavation and backfilling activities are also needed to replace the native soil of the site if it has any problem such as swelling or organic materials. When a large quantity of excavation or backfilling is involved in a project it increases its total cost especially if the native soil is rocky or hard to excavate.

The unit of this factor will be expressed during the analysis stage by stating the needed quantity in (m^3) of backfill or excavated materials.

3.2.3.12 Number of Classrooms

The number of classrooms is an important factor that determines the size of the school and accordingly its cost. As the number of classrooms increases, the number of students, number of floors and all associated supporting facilities such as laboratories and toilets also increase. This increase will clearly have a great impact on the budget needed to build the school project.

3.2.4 Data Collection

1. *Investigating the standard designs available to the Ministry of Education.*

The purpose of this investigation is to find out the type of construction and designs that the Ministry of Education uses in its school building projects. With the knowledge of the different components of the schools we can assess the basic differences between the models.

2. *Studying the Old Files and Records of the School Projects.*

The purpose of this process is to collect all data needed for developing a valid and reliable cost estimate models for the Ministry of Education. These data were collected and filled into a form (Table 1 in Appendix A). Basically, there was one form that had to be filled for every school selected in our sample. Each form presents all data required for every school, such as area, number of classrooms, bid price, etc. Then the data for all schools were filled in another table (Table 2 in Appendix A) which made it easier for the purpose of conducting the analysis.

3.2.5 Sample Selection

It is known that the construction boom (1975-1981) in Saudi Arabia affected all components of the construction industry. It inflated the prices of construction materials, equipment and labor. Exaggerated and inflated profit margins were also not uncommon.

After 1981, the construction boom ended and the construction industry, like other sectors of the Saudi economy, started to operate more conservatively to accommodate the government's new conservative attitude towards spending.

Since all governmental ministries are still operating under this conservative attitude, our sample will be selected so that it will focus on those schools built after the construction boom period. This is to keep the homogeneity between our selected sample and the current situation of the construction industry and the economic condition of the Saudi Government. Accordingly, those schools which were awarded during 1994 were selected due to the availability and easy access to all required information needed for completing this research.

CHAPTER FOUR

CONSTRUCTION COST PROBABILITY ASSESSMENT MODEL

This chapter presents the characteristics of schools built by the Ministry of Education for boys, the method that is currently used to estimate the cost of school construction and the process for developing statistical models for estimating the construction cost of boys' school building.

4.1 CHARACTERISTICS OF BOY SCHOOLS

The Ministry of Education builds boy schools using seven design models that have been in use for the last two decades. These design models are designated by certain code numbers, namely 2,3,4,5,6,10 and 11. One design model might be used for building an elementary, an intermediate or a secondary school depending on the need and the limitation of the design itself. The different school models and their use are illustrated in Table 2.

Table 2: Types of School Design Models.

Design Model	# of Classrooms	Elementary Level	Intermediate Level	Secondary Level
2	18	Yes	Yes	No
2	28	Yes	Yes	No
2	26	No	No	Yes
3	12	Yes	Yes	Yes
3	20	No	No	Yes
3	22	Yes	Yes	No
4	9	Yes	Yes	Yes
5	14	No	Yes	No
5	19	Yes	No	No
5	21	No	Yes	No
5	28	Yes	No	No
6	18	Yes	Yes	Yes
6	24	Yes	Yes	Yes
6	30	Yes	Yes	Yes
10	21	No	Yes	No
11	15	Yes	Yes	Yes
11	20	Yes	Yes	Yes
11	25	Yes	Yes	Yes
11	30	Yes	Yes	Yes

Also, one design model might be used for building schools with a different number of floors and a different number of classrooms. The minimum and maximum number of floors and number of classrooms are shown in Table 3.

Table 3: Maximum and Minimum Floors and Classrooms

For all Design Models

	Minimum	Maximum
Number of Floors	1	4
Number of Classrooms	9	30

It may be noticed from Table 1 presented in Appendix (C) that school design models number 4, 6 and 10 have very few data points to analyze. It was for this reason only that the scope of this research was limited to school design models number 2,3,5 and 11. The description of these models as well as their components and characteristics are discussed in the following sections.

4.1.1 School Design Model # 2

This school model is designed in such a way that it can be used for elementary, intermediate or secondary levels. The basic design concept is the same for all school levels but the difference in use might be achieved by changing the number of classrooms or the number of floors. The main characteristics for the school project regardless of its level are as follows:

1. Administration

A- Principal room

B- Social Advisor Room

C- Secretary room

2- Classrooms

A- The number of class rooms could be 18, 26 or 28.

B- 2 teacher conference rooms

3- Number of Floors

A- The school could be two storey or three storey building.

4- Technical Laboratories

A- 2 laboratories for Intermediate Level

B- 4 laboratories for Secondary Level

5- Activities:

A- Library

B- Gymnasium

C- Multipurpose room

6- **Services**

A- Stores

B- Toilets

C- Snack Bar

D- Mosque

The concrete area for this school design model ranges between 2553 m^2 and 3720 m^2 depending on the level as well as the number of classrooms. All classrooms and supporting facilities are distributed on the periphery of the building while the core is left to be used as the students' yard.

4.1.2 School Design Model # 3

This school model is designed in such a way that it can be used for elementary, intermediate or secondary levels. The basic design concept is the same for all school levels but the difference in use might be achieved by changing the number of classrooms or the number of floors. The main characteristics for the school project regardless of its level are as follows:

1. Administration

A- Principal room

B- Social Advisor Room

C- Secretary room

2- Classrooms

A- The number of class rooms could be 12, 20 or 22.

B- 2 teacher conference rooms

3- Number of Floors

A- The schools could be two storey or three storey building.

4- Technical Laboratories

A- 2 laboratories for Intermediate school

B- 4 laboratories for Secondary school

5- Activities:

A- Library

B- Gymnasium

C- Multipurpose room

6- Services

A- Stores

B- Toilets

C- Snack Bar

D- Mosque

The concrete area for this school design model ranges between 2221 m^2 and 3220 m^2 depending on the level as well as the number of classrooms. All classrooms and supporting facilities are distributed on the periphery of the building, while the core is left to be used as the students' yard.

4.1.3 School Design Model # 5

In contrast to all other design models, this school model is designed in such a way that it can be used for elementary and intermediate school levels only. The basic design concept is the same for both school levels but the difference in use might be achieved by changing the number of classrooms or the number of floors. The main characteristics for the school project regardless of its level are as follows:

1. Administration

A- Principal room

B- Store room

C- Secretary room

D- Archive

E- Ambulance room

2- Classrooms

A- The number of class rooms could be 14, 19, 21 or 28.

B- 1 Teacher conference room

3- Number of Floors

A- The schools could be three storey or four storey building.

4- Technical Laboratories

A- 2 laboratories for Intermediate level

5- Activities:

A- Library

B- Drafting room

C- Multipurpose room

D- Audio Room

6- Services

A- Stores

B- Toilets

C- Snack Bar

The concrete area for this school design model ranges between 2536 m^2 and 3453m^2 depending on the level as well as the number of classrooms. All classrooms and supporting facilities are distributed on 3 sides of the building only, while the core is left to be used as the students' yard.

4.1.4 School Design Model # 11

In contrast to other design models available with the Ministry of Education, this model is distinguished by its large area. This large area might be utilized in building more classrooms, activity areas or service areas. The basic design concept is the same for all school levels but the difference in use might be achieved by changing the number of classrooms only. The number of floors is two regardless of whether this design model will be used for elementary, intermediate or secondary levels. The main characteristics for the school project regardless of its level are as follows:

1. Administration

A- Principal room

B- Store room

C- 2 Principal Assistance rooms

D- 1 Conference room

E- Store room

F- 2 Monitoring rooms

2- Classrooms

A- The number of class rooms could be 15, 20, 25 or 30.

B- 2 teacher conference rooms

C- Physical Education Advisor room

D- Drafting room

E- Computer room

3- Number of Floors

A- The school is two storey building

4- Technical Laboratories

A- 2 laboratories for Intermediate level

B- 4 laboratories for secondary level

5- English Laboratory

A- The school has one English Laboratory.

6- Computer Lab

A- The school has one computer laboratory.

7- Activities:

A- Library

B- Gymnasium (might be included or not)

8- Services

A- Stores

B- Toilets

C- Snack Bar

D- Dressing room

The concrete area for this school design model ranges between 5600m^2 and 6350m^2 depending on the level and the number of classrooms.

4.2 EXISTING COST ESTIMATE SYSTEM

Regarding the already existing cost estimate system in the Ministry of Education, an open ended questionnaire was developed (Appendix A) in order to identify the historical background, the working system and finally to know the advantages and disadvantages of using such a system from the point of view of the end user. It was intended to include all the employees who are involved in the estimating process in the Project Studying Department, but the researcher was restricted to interview the Head of the Department only. It was assumed that the Head of the Department possessed the necessary and acceptable information to serve the purpose of this study and to reach a conclusion about the current used estimating system. Therefore, the questionnaire was filled-in through personal interview with the Head of the Projects Studying Department. The results that were obtained from this questionnaire are divided into three main parts: historical background, explanation of the used system and finally stating the advantages and disadvantages of using such a system from the end user's point of view.

4.2.1 Description of the Used Estimating System

The respondent indicated that in 1990, the Project Studying Department developed the current estimating system. Since then, the Ministry has been using this system to prepare budget estimates for their school projects. He also mentioned that the estimating process is conducted by engineers.

The respondent was asked several questions seeking detailed information regarding the nature of the existing estimating system, the required input parameters for appling the system (inputs), the system estimating procedures, the produced end results (output) of the system, the use of the produced estimate, and the process that the Ministry of Education follows for having the approval on the needed budget from the Ministry of Finance.

The respondent indicated that the estimating system is simple and purely manual. Although the system is manual, there is no standardized form to be filled or to be used for calculating the cost of the project. This system is based on using the concept of floor area unit estimate in which the cost of the project is determined based on a fixed price of SR per concrete square meter. This fixed unit price is established based on the Ministry's experience acquired from all previous projects and continuously updated whenever it is found that it is no longer valid.

Also, this unit price is applied for preparing estimates for schools regardless of the design model. So, determining the estimated construction cost for any school project involves calculating the total concrete area of the school and then multiplying it by a predetermined unit price. This estimating system is simple but it does not provide the breakdown cost for every important item of the project, such as the concrete work, earth work, finishing, etc. But instead it gives a lump sum cost for the whole project.

After obtaining the approximate cost of the school project using the above described method, it is submitted to the bid awarding committee for approval. The respondent was also asked about the criteria which this committee based their approval on and the necessary steps followed to get the approval of the Ministry of Finance on the estimated budget prepared by the Ministry of Education, but unfortunately he was very conservative since he considered such information as confidential for both ministries. The organizational chart was also requested by the researcher in order to complete his overall understanding of the execution of the estimating process but the respondent refused for the same reason as mentioned above.

4.2.2 Strengths and Weaknesses of the Used Estimating System

It was found that the current system has weaknesses exceeding its strengths by far. This low satisfaction with the system may result from users comparing the available system with some other system that has been introduced to them somewhere else. The only strength reported for the system is that it produces estimates in a short time (one day).

On the other hand, the system was not found to take care of the different characteristics of the school projects, such as:

1. the different available design models
2. different available soil types
3. different foundation systems used
4. different effort spent on site leveling activities such as backfilling and excavation.

From the above discussion and evaluation, it can be noticed that the Ministry of Education is depending on a single factor (concrete area) for making its budget estimate. Since involving more variables in the budget estimating process might make the system more flexible in predicting the cost of different projects with different influencing factors, it is believed that involving some factors related to the site conditions of the project or the level of school itself in the estimating process might increase the accuracy of the system and make it more flexible and reliable.

Moreover, if a more advanced cost estimate system is developed, it will be more convincing to the Ministry of Finance and it could serve as a common base for budget estimating for both the Ministry of Education and the Ministry of Finance.

Based on the above discussion, the researcher found it necessary to develop several alternative statistical models for estimating the budgetary prices for the construction cost of school projects in which most of the problems faced in the existing system could be eliminated or minimized.

4.3 DEVELOPMENT OF CONSTRUCTION COST ESTIMATE MODELS

This part describes the development of models to assess the probability distribution of the construction cost of boys schools in which some general information about the characteristics of a school project are known at early stage of the design process. The following sections present the prediction models in which they will be used to statistically predict the construction cost for schools. Finally the validity and the predictability of the developed models will be measured.

4.3.1 Model Development

The probability distribution of the construction cost of a boys school, j, CC_j, given the knowledge of physical (Pj) characteristics is normal with a mean $\sum_{k=1}^{n} \beta_{jk} \; X_{jk}$ and a variance of σ_E^2

$$\text{In notation } \{CC_j | P_j\} \approx N\left(\sum_{k=1}^{n} \beta_{jk} X_{jk}, \sigma_E^2\right)$$

The mean of the distribution is estimated by a regression equation. The variance is taken as the mean-squares of the residuals of the regression equation.

A sample of bid prices for 170 boys' schools projects that were submitted to the Ministry of Education and awarded to different Saudi Contractors in 1994 was used to determine:

(1) which of the many potential physical factors are to be included in the construction cost regression equation;

(2) what type of mathematical combinations govern the construction cost and those potential factors Pj;

(3) the regression equation factors' weights, β_k, and

(4) the variance of estimates σ_E^2.

These objectives were achieved through an intensive graphical and statistical analysis in the mainframe using the **SAS** package. Then the normal linear regression method was used to calculate the factors' weights and the standard error of estimates.

4.3.1.1 Determination of the Construction Cost Regression Equation

As discussed before in section 3.2.2.2 the potential physical factors which were found to be responsible for the variations in the construction costs of boys schools are as follows:

1. Number of classrooms
2. Length of fence
3. Project's execution time
4. Footing type
5. Soil bearing capacity
6. Soil improvement

7. Excavation material

8. Backfilling material

9. Level of school

10. Concrete area of school

11. Number of laboratories available in the school

12. Length of retaining walls

Observations on the above factors and bid prices for 192 schools were collected from the archive of the Ministry of Education. Out of these 192 data points, 170 school projects were utilized to conduct the statistical analysis and the remaining were randomly selected to test and check the predictability of the models developed . The bid prices for all the schools in the sample were submitted to the Ministry of Education in 1994 by different Saudi contractors.

The relationships between the cost (dependent variable) and the various independent variables such as the number of classrooms, soil improvement, execution time, etc., were examined and scatter plots were produced.

The scatter plots were visually studied very carefully but, unfortunately, the relationships between the cost of the project and the various independent variables were not clearly identified.

Since no clear relationship was observed between the cost of the project and the various independent variables, observations on those independent variables were transformed mathematically into logarithms, exponential, inverse, squares, etc. The cost of the project was again plotted against the transformed data of the independent variables and scatter plots were produced and studied, but again no definite relationships were concluded.

Since the relationships between the different factors and the construction cost of school project were not clearly identified by the produced scatter plots and since the linearity was proven to be a good assumption in similar studies (Bozay, 1981 & Mashoor, 1993), it was decided to assume linear relationships between the various independent variables and the construction cost of school project. This assumption will be tested and verified after developing the models.

4.3.1.2 Determining the Probability Distribution Parameters

Since it was concluded from the graphical analysis that no clear relationship exists between the cost of the project and any of the independent variables, it was decided to develop the best linear regression function for the collected data. The general form of the potential regression equation is as follows:

Predicted Cost = $\beta_0 + \beta_1 (X_1) + \beta_2 (X_2) + \beta_3 (X_3)$ +------ $\beta n (X_n)$

where X_1, X_2, X_3 and X_n are the independent variables such as concrete area, number of class rooms, bearing capacity of soil, etc.

The following assumptions are made:

1. The construction cost is a linear function of the independent variables.
2. No relation exists between the independent variables.

The validation of the first assumption was checked by analyzing and studying the residual plot, while the validation of the second assumption was performed by studying the correlation matrix of all independent variables shown in Appendix B. The linear least squares regression was performed using the **SAS** package to calculate the regression coefficients, β_k, and the variance of estimates, σ_E^2, of the construction cost regression equation.

A correlation factor of (R = 1) implies a very high association between two variables and a correlation factor (R = 0) means no correlation between those two variables. Since the |R| value of one and zero between two variables represents the two extremes of being correlated or non-correlated, it was decided to consider 0.5 correlation factor as the cut off point since it can be considered as a value unbiased to any of these two extremes. So, it was decided to eliminate one independent variable if it has a correlation factor of ($|R| \geq 0.5$) with another variable. Accordingly, when two independent variables were found with a correlation factor ($R \geq 0.5$), different statistical models were produced. Each model contains in addition to the other different factors one of the two correlated variables. The set of variables which produces the model with the highest R^2 were selected.

After examining the developed correlation matrix, the variables found with a high correlation factor ($|R| \geq 0.5$) are as follows:

Execution time	-	Number of classrooms
Footing type	-	Number of class rooms
Quantity of back filling material	-	Length of fence
Concrete area	-	Number of class rooms
Concrete area	-	Length of fence
Concrete area	-	Execution time
Level of schools	-	Number of technical laboratories
Quantity of excavated material	-	Length of retaining walls
Quantity of back filling material	-	Length of retaining walls

The different regression models were developed to investigate which set of non-related independent variables produces the needed model with higher R^2. This analysis indicated that the following variables produce the highest R^2:

1. Soil improvement
2. Concrete area
3. Footing type
4. Soil bearing capacity
5. Number of technical laboratories
6. Quantity of back filling material
7. Quantity of excavated material

Therefore the potential form for the general regression equation for the construction cost of school projects is as follows:

CC = β_0 + β_1 (soil improvement) + β_2 (concrete area) + β_3 (footing type) + β_4 (soil bearing capacity) + β_5 (Number of technical laboratories) + β_6 (Quantity of back filling material) + β_7 (Quantity of excavated material).

Where:

CC	=	School Construction Cost
β_0	=	Regression Constant
β_k	=	Factor weights, K = 1, ---------7

The linear least squares regression was performed using the step wise procedure available in the **SAS** package to calculate the regression coefficients, **βk,** and the variance of estimates, σ_E^2, of the construction cost regression equation (table 4). The regression analysis correlates the 170 schools' construction cost to the five potential factors. The R^2 value for the regression is 90.04 percent. The five factors explain 90.04 percent of the variation in the schools' construction cost.

Since the confidence level for the selected model represented by its significance level is considered very high (99.99%), it can be concluded that the model is statistically significant. Regarding the parameters themselves, it can be seen that their confidence level is also very high (97.48% - 99.99%) and we can conclude that all the above parameters will be considered in the selected linear statistical model.

TABLE 4 : REGRESSION COEFFICIENTS AND THEIR SIGNIFICANCE LEVEL FOR THE GENERAL REGRESSION MODEL

Regression Equation :-

$CC = \beta_0 + \beta_1$ (Concrete Area) + β_2 (Quantity of excavated mateial) + β_3 (Footing type) + β_4 (Quantity of backfilling material) + β_5 (# of Technical laboratories)

Coefficient of Multiple Determination $R^2 = 0.9004$

Multiple Correlation Coefficient R = 0.9488

Source Of Variation	Degree of Freedom	Sum Of Squares	Mean Squares	F Ratio	Significant At
Regression	5	1.5527680376E15	310553607519889	296.55	99.99%
Residual	164	171746075211259	1047232165922	-	-

Variable	Regression Coefficient	Standard Error of Estimate σ_E	Variance of Estimate σ_E2	T-Ratio	Significant At
Constant	-1257537.45	443313.17	1.965265667E11	2.83	99.49%
Concrete Area	1721.36	63.10	3981.61	27.27	99.99%
Quantity of Excavated Material	55.0257	24.36	593.4	2.25	97.48%
Footing Type	1261193.45	372008.5	1.383903241E11	3.38	99.91%
Quantity of Backfilling Material	71.0794	14.515	210.685	4.89	99.99%
# of Technical Lab.	355491.708	51747.2	2677772708	6.86	99.99%

Consequently, the liner general statistical model for all school design models built by the Ministry of Education in Saudi Arabia is as follows:

Predicted Cost: **-1257537.45 + 1721.36 (Conc. area) + 55.025 (Excav) + 1261193.45 (Foot type) + 71.079 (Fill) + 355491.71 (Lab No.)**

where

Conc. area	=	Concrete area (m^2)
Excav	=	Quantity of excavated materials (m^3)
Foot type	=	Footing type
Fill	=	Quantity of backfilling materials (m^3)
Lab No.	=	Number of technical laboratories

As can be seen from the above statistical model, five different independent variables significantly contribute to the cost of all boys school models built by the Ministry of Education. These factors can be grouped into two categories. First, the substructure category including excavation, backfilling and footing type. Second, the superstructure category including concrete area and the technical laboratories. Having both the categories included in the developed model indicates that they are equally important in predicting the cost of boys schools regardless of their design models.

The design models are basically different in their concrete area and accordingly the needed footing type. So, it was expected to find those two variables significantly contributing to the developed general model. It is known that as the concrete area increases, the footing area also increases provided that the bearing capacity is kept almost the same. The isolated footing area might increase until it exceeds 60% of the total area of the project where it becomes advisable to use mat foundation instead of isolated footing. The construction of mat foundation has a great impact on the overall cost of the project since it increases the amount of used materials, manpower and machinery. The effect of increasing the concrete area by $100m^2$ and shifting from isolated footing to mat foundation will produce an increase in the construction cost of 2.1% and 1.8% respectively.

The technical laboratories in a school building also play an important role in determining the total cost of a project. This is due to the sophisticated machines, instruments and expensive apparatus usually involved in these technical laboratories. The high cost areas that are included in technical laboratories are acid resistant tiles for floors and walls, smoke and fire detectors, manual warning machines, lighting warning systems and gas leak warning systems. The cost of all these additional requirements has a great impact on the total cost of the project and accordingly it makes a lot of difference if the school will have labs or not. This great impact was found to have 10% increase on the total cost of the project for every two additional laboratories.

One of the most variable aspects of construction projects is the site conditions. In the case of the school projects built by the Ministry of Education, the superstructures are always identical for the same design model but some differences might exist in the site conditions allocated to the selected project. In order to prepare the site for the new project, some activities such as excavation or backfilling might be involved. These activities are usually needed to level the site or to put it on a certain level based on the project specification and recommendation. It might be expected that excavation and backfilling activities would have less effect than other parameters on the total construction costs of school project since they have lower unit costs compared to the other parameters. It was found that the effect of increasing the quantity of excavation and backfilling materials by $100m^3$ each will produce an increase in the total construction cost of 0.075% and 0.1% respectively.

It may be noticed that some of the potential variables which were thought to have an effect on the total construction cost of the school project such as soil improvement and soil bearing capacity are not included in the developed linear statistical model. In the case of soil bearing capacity, it can be seen from the collected data that the bearing capacity of the majority of the selected sites (95%) ranges between 1.0 Kg/cm^2 and 2.0 Kg/cm^2. It was found from documents seen in the archive of the Ministry of Education that the effect of increasing the bearing capacity on the square area of this isolated footing as well as the increase in steel reinforcement ranges between 20% to 25%. Since this increase in the raw materials cost is considered negligible compared to the total construction cost of the project, it was found that this factor does not have a great impact on the total construction cost of the school project, and is therefore excluded from the developed model.

It is known that if the soil of a project's site is to be improved, it will have an impact on the overall construction cost for that project. This impact might be high or low depending on the method or techniques utilized to do the job. It was found from the collected data that only 10% of projects needed soil improvement for their sites. Moreover, the soil was improved in 60% of them by inexpensive methods such as excavation, backfilling and compaction, and accordingly it did not add much to the total cost of the project. All these factors render the cost of applying soil improvement to the school project and make them less important than the other independent variables represented in the developed statistical equation.

Table 5: The Average Effect of Increase in one of the Parameters on the Overall Construction Cost of the School Project for the the general regression model.

Parameter	Amount of Increase	Percentage of Increase
Concrete Area	$100m^2$	2.1%
Quantity of Excavation	$100m^3$	0.075%
Footing type	isolated to mat	1.8%
Quantity of backfilling	$100m^3$	0.1%
Number of technical laboratories	2	10%

4.3.1.3 Testing of Assumption

When a regression model is applied in practice, one cannot usually be sure in advance that the model implied is appropriate for the situation in hand. Consequently, the model considered needs to be checked for suitability and its agreement with the earlier assumptions made to develop it. As stated earlier, the following assumption were made:

i The regression function is linear

ii The distribution is normal

iii There is a constant variance over all schools construction cost.

In the following paragraphs, the above mentioned assumptions will be verified and checked.

4.3.1.3.1 Linearity Assumption

In order to test the developed model for its aptness and linearity, the residual plot was analyzed and studied. This graph was obtained by plotting the residuals against the fitted values (Figure 9). If the specified regression model is linear, the residuals will tend to scatter at random around the zero line when plotted against the fitted values. As can be seen in Figure 9 there is no pattern of systematic departure and all points are scattered normally around the zero line. This reveals that the selected statistical general model for those schools built by the Ministry of Education is linear and apt.

4.3.1.3.2 Normality and Constant Variance Assumptions

In order to check the normality assumption, and to be sure that a constant variance exists over all schools construction cost, the lack of normality test was performed. In this test, the standardized residuals were calculated. If the distribution was normal with the constant variance and the sample size is considerably large (greater than 30 observations), the standardized residuals will tend to follow the standard normal distribution (figure 10). Hence, about half of the standardized residuals should be positive and half negative, About 68% of the standardized residual distribution should fall between -1 and +1 and at least 95% of the area under the normal distribution should be contained within t (0.025: n-2) and t (0.975: n-2) (i.e. $\mu \pm 2\sigma$) (Netir, Warerman and Whitemore, 1978).

In order to check if the developed model complies with the above mentioned criteria of normal distribution, the standardized residuals were calculated and listed in table 6.

Where

Cost	=	the bid price for each observation point.
Predicted	=	the predicted cost using the developed model
Residual	=	cost - predicted
Standardized Residual	=	$\frac{\text{Residual}}{\sqrt{\text{MSE}}}$
MSE	=	mean square error

Based on the analysis that was conducted on the results listed in table 6, the following was found

$$t \quad (0.025 : 168) \quad = \quad -1.96$$

$$t \quad (0.975 : 168) \quad = \quad +1.96$$

$$-1 < 77\ \%\ \text{of the distribution} < +1$$

54% of the distribution are negative

-1.96 < 97% of distribution < 1.96

Since the above analysis complies with the different test criteria, then it can be concluded that this analysis provides no evidence of any departure form normality and the variance is constant over all school construction cost.

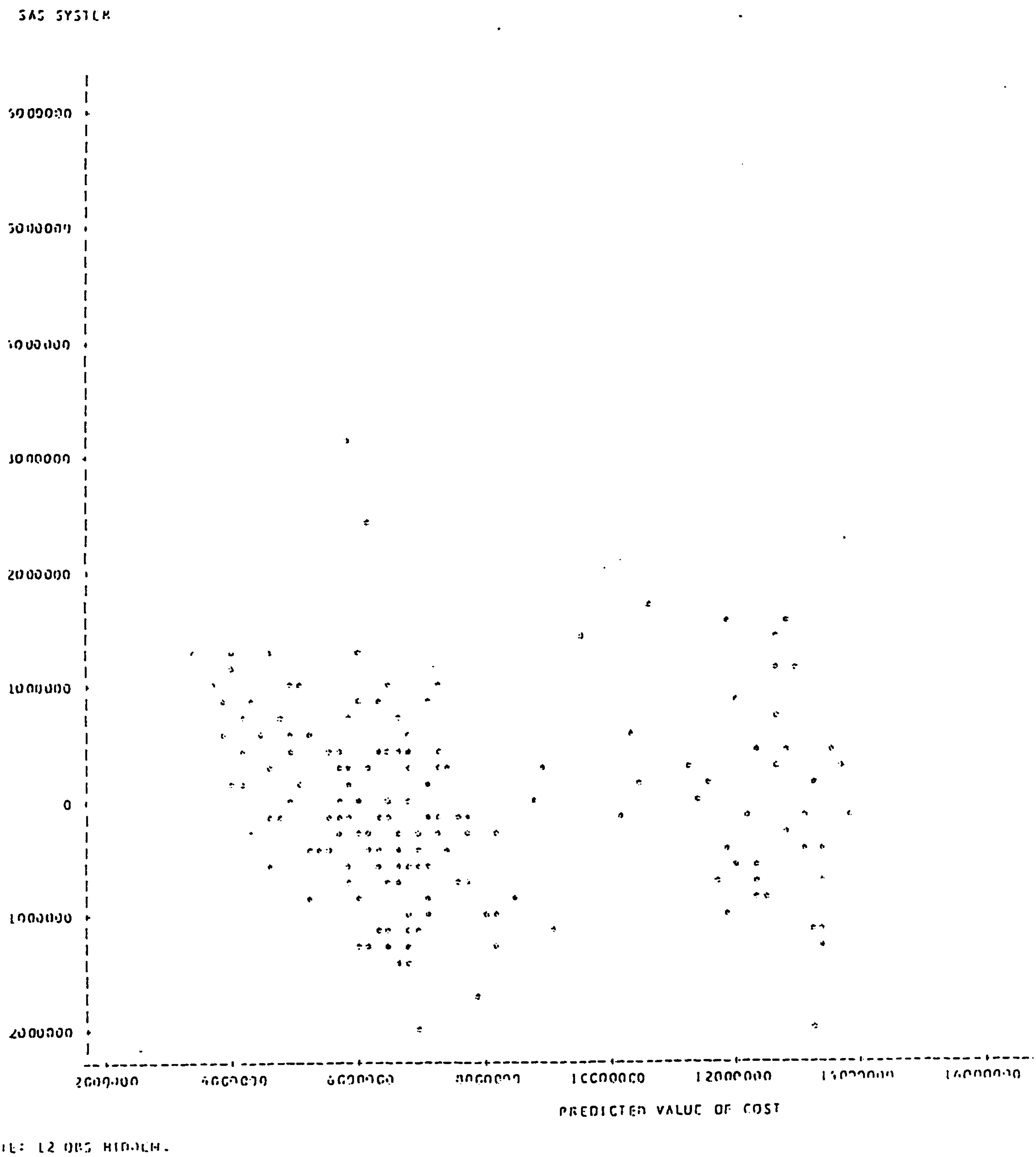

Figure 9: Residual Plot For the General Model

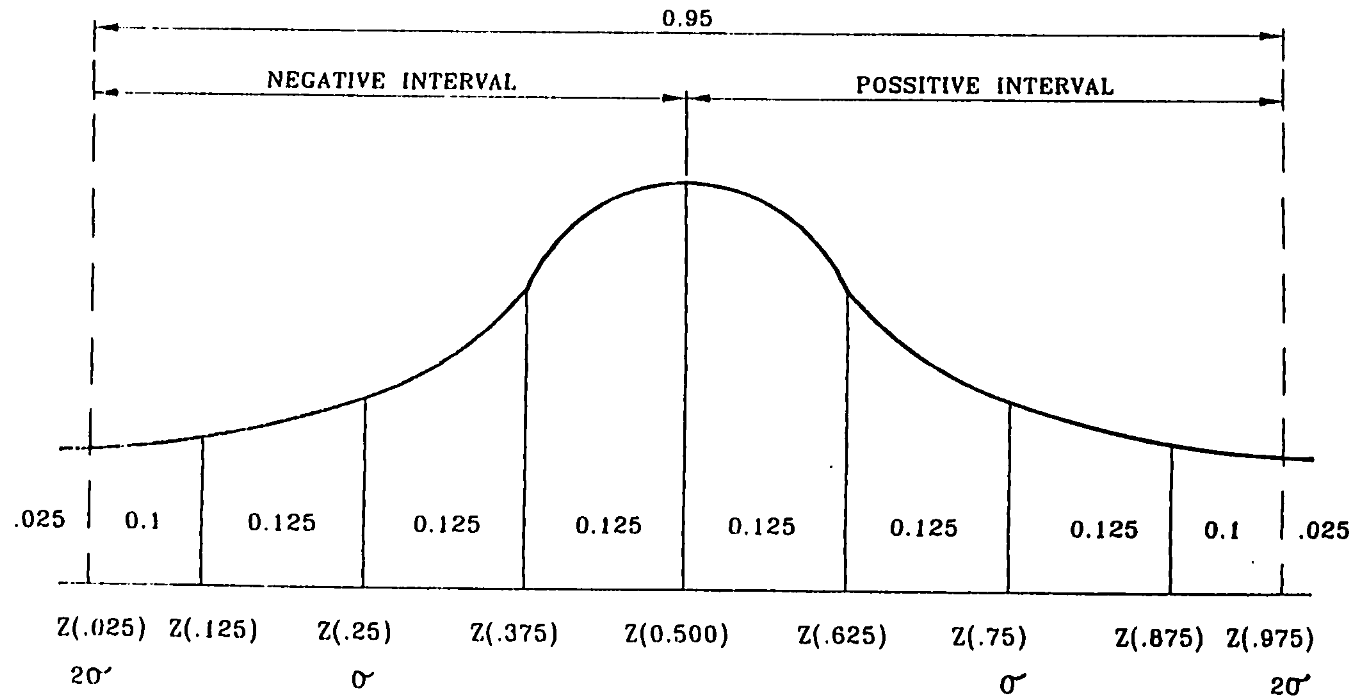

FIG:10. PARTITIONING OF A NORMAL DISTRIBUTION CURVE INTO 8 EQUAL INTERVALS FOR LARGE SAMPLE SIZE.

Table 6: : RESIDUAL TABLE FOR THE GENERAL MODEL

NO	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
1	5553587	5706523	-152936	-0.149447361
2	6904694	7685788	-781094	-0.763276381
3	5436815	4490999	945816	0.92424089
4	5733160	4770890	962270	0.940319556
5	6374790	6416515	-41725	-0.040773207
6	6901734	7167438	-265704	-0.259642998
7	6559178	6886191	-327013	-0.319553472
8	5073788	4408815	664973	0.64980423
9	7176342	6231606	944736	0.923185526
10	4961254	6530360	-1569106	-1.533312956
11	11543033	12323307	-780274	-0.762475087
12	11411937	11881696	-469759	-0.459043277
13	13490045	11338668	2151377	2.10230171
14	12356315	10844538	1511777	1.477291694
15	6791200	7470782	-679582	-0.664079983
16	6392235	5885607	506628	0.495071255
17	5874050	6631858	-757808	-0.740521561
18	11384782	11527250	-142468	-0.139218147
19	5868298	5598645	269653	0.263501917
20	5329678	6468723	-1139045	-1.113062123

Table 6: Continued

NO	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
21	8742473	8417651	324822	0.317412451
22	7883848	7292070	591778	0.578278889
23	7229710	7463215	-233505	-0.228178493
24	5271682	4433832	837850	0.818737714
25	4863248	3954743	908505	0.887780996
26	7268715	6638350	630365	0.615985677
27	12935335	12444823	490512	0.479322878
28	12914745	12392322	522423	0.510505953
29	11421496	12547396	-1125900	-1.100216975
30	13107765	12404003	703762	0.68770841
31	5833170	6579572	-746402	-0.729375744
32	11339652	11194544	145108	0.141797926
33	5950785	5380727	570058	0.557054346
34	6337282	5639353	697929	0.682008467
35	7007962	7144690	-136728	-0.133609083
36	6941803	7181991	-240188	-0.234709046
37	6819579	6542095	277484	0.271154283
38	6617453	7128222	-510769	-0.499117794
39	7538941	7241949	296992	0.290217284
40	7696680	7807584	-110904	-0.108374157
41	5955973	5720833	235140	0.229776196
42	5428406	4735673	692733	0.676930994

Table 6: Continued

NO	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
43	7698578	7344244	354334	0.346251249
44	7655002	7334580	320422	0.31311282
45	6395654	6543578	-147924	-0.14454969
46	6594874	6885192	-290318	-0.283695525
47	11609838	12605633	-995795	-0.973079814
48	5076755	5775189	-698434	-0.682501948
49	5862659	6579699	-717040	-0.700683524
50	6134082	7827131	-1693049	-1.654428679
51	12327863	13508590	-1180727	-1.15379331
52	5985308	6702467	-717159	-0.70079981
53	11252936	12486746	-1233810	-1.205665429
54	4453455	5085012	-631557	-0.617150486
55	4343296	4581267	-237971	-0.232542618
56	5464115	6416514	-952399	-0.930673725
57	8307439	7715874	591565	0.578070748
58	6832657	8009885	-1177228	-1.150374127
59	6100363	6127284	-26921	-0.026306902
60	13956005	13264485	691520	0.675745664
61	12717476	13493745	-776269	-0.758561445
62	14075180	12944717	1130463	1.104675888
63	9125940	8295919	830021	0.811087302
64	12869200	11993469	875731	0.855754607

Table 6: Continued

NO	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
65	5890731	4784723	1106008	1.080778734
66	7440296	7611109	-170813	-0.166916566
67	7935874	7332515	603359	0.589595714
68	5895825	4629761	1266064	1.237183679
69	12462787	12697088	-234301	-0.228956335
70	7655616	9105560	-1449944	-1.416869173
71	5660981	5643081	17900	0.017491681
72	11018318	11414469	-396151	-0.387114357
73	4134785	4378092	-243307	-0.237756898
74	10920579	11561179	-640600	-0.625987205
75	10872666	10729405	143261	0.139993058
76	11961795	12540714	-578919	-0.565713217
77	13777792	12271476	1506316	1.471955265
78	13410720	12631684	779036	0.761265327
79	4825280	5992006	-1166726	-1.140111689
80	5758779	6504263	-745484	-0.728478685
81	5093569	5987008	-893439	-0.873058668
82	5718324	6542770	-824446	-0.805639474
83	4780251	5983843	-1203592	-1.176136735
84	5723083	5411561	311522	0.304415838
85	6299232	7049446	-750214	-0.733100788
86	4434368	4532518	-98150	-0.09591109

Table 6: Continued

NO	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
87	15283084	16437371	-1154287	-1.127956436
88	14107256	13195443	911813	0.891013536
89	12569974	12700671	-130697	-0.127715657
90	12071516	13687065	-1615549	-1.578696539
91	12048217	13092732	-1044515	-1.020688457
92	12882354	13716681	-834327	-0.815295078
93	13918823	13105343	813480	0.794923621
94	14398474	13643509	754965	0.737743413
95	12758486	12509443	249043	0.243362054
96	5885340	5503051	382289	0.373568565
97	5189223	6452307	-1263084	-1.234271656
98	5686894	5433577	253317	0.247538559
99	4813824	4883001	-69177	-0.067598996
100	4093342	3891463	201879	0.197273917
101	7821456	8161144	-339688	-0.331939341
102	5688770	6022870	-334100	-0.326478809
103	6181035	6657881	-476846	-0.465968615
104	5098395	4405721	692674	0.67687334
105	5431683	5507618	-75935	-0.074202839
106	6095409	5358252	737157	0.720341633
107	7344552	6200417	1144135	1.118036014
108	5538262	6730898	-1192636	-1.165430653

Table 6: Continued

NO	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
109	4938182	4421769	516413	0.504633048
110	6130537	6456224	-325687	-0.318257719
111	6931120	7497332	-566212	-0.553296078
112	8541311	5997509	2543802	2.485775061
113	4673077	4755071	-81994	-0.080123626
114	5437666	4939189	498477	0.487106188
115	6083932	4881684	1202248	1.174823393
116	6045409	5537318	508091	0.496500882
117	6529180	5374997	1154183	1.127854808
118	4942961	4215894	727067	0.710481797
119	6960810	7362241	-401431	-0.392273915
120	6390231	7197585	-807354	-0.788937362
121	7337905	6595061	742844	0.725898906
122	8335448	6791549	1543899	1.508680955
123	6908921	6412430	496491	0.485165491
124	5275020	4451922	823098	0.804322224
125	6957923	6772273	185650	0.181415118
126	10990783	8425344	2565439	2.506918497
127	7577367	7560347	17020	0.016631755
128	8850205	5868374	2981831	2.913812135
129	7143186	7836818	-693632	-0.677809486
130	7408094	7172475	235619	0.23024427

Table 6: Continued

NO	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
131	5118840	4795878	322962	0.315594879
132	9995390	10089030	-93640	-0.091503968
133	4350750	4565335	-214585	-0.209690079
134	11810756	12624349	-813593	-0.795034043
135	10643880	10067247	576633	0.563479363
136	11684579	11799716	-115137	-0.112510598
137	5932772	5419148	513624	0.501907668
138	7078125	7264805	-186680	-0.182421623
139	6156745	6749355	-592610	-0.57909191
140	6749661	6750728	-1067	-0.001042661
141	5203988	5518257	-314269	-0.307100176
142	6873643	7045148	-171505	-0.167592781
143	5059463	5670861	-611398	-0.597451335
144	6008365	5789515	218850	0.213857789
145	4587976	4885536	-297560	-0.290772327
146	4761702	6417130	-1655428	-1.617665855
147	5251840	6453245	-1201405	-1.173999622
148	5342725	7294207	-1951482	-1.906966536
149	5103970	6451640	-1347670	-1.316928156
150	5695288	6599474	-904186	-0.883560517
151	5463114	7202147	-1739033	-1.699363733
152	5991870	7234360	-1242490	-1.214147428

Table 6: Continued

NO	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
153	5953221	6463496	-510275	-0.498635063
154	6935347	8480994	-1545647	-1.510389081
155	4036257	4500904	-464647	-0.454047888
156	5297435	6154757	-857322	-0.837765536
157	5986049	5749853	236196	0.230808108
160	6259610	6711999	-452389	-0.442069506
161	4645076	3949016	696060	0.680182101
162	6007055	6510514	-503459	-0.491974543
163	4915859	4468546	447313	0.437109295
164	4724756	3937322	787434	0.769471759
165	7389708	7160330	229378	0.224145634
166	7036454	6619247	417207	0.407690047
167	7386606	6160952	1225654	1.197695476
168	13722367	14977766	-1255399	-1.22676196
169	13288946	14098651	-809705	-0.791234733
170	11562856	11330984	231872	0.226582743

4.3.1.4 Validity of the Model

In order to test or to verify the validity of the model, ten schools were selected randomly, excluded from the data base and withheld for checking the predictability of the developed model. The reliability of the model's predictability was examined by comparing the bid price with the model estimates (Table 7). As can be seen from Table 7 that the developed model tends to under estimate the construction cost of the project. On the other hand, the model possesses a balanced behavior when it over or under estimates the project costs (i.e. the percent difference between bid price and predicted cost reached almost the same value with positive and negative signs). This balanced action adds more credits to the reliability and the predictability of the developed statistical model.

According to the AACE, the accuracy range for the conceptual estimate is defined by -30%, + 30% (Mashhoor, 1993). Since this model is primarily designed for the purpose of preliminary estimating and moreover, the variations in the predicted and bid price costs falls within the acceptable limits (-30%, +30%), then it can be concluded that the developed model is the most representative statistical model of the construction costs of boys' schools built as per design models number 2, 3, 5 and 11.

Table 7: A Comparison Between the Bid Price and the Predicted Cost for the general regression model.

Project Name	Design Model	Predicted Cost	Bid Price	Percent Difference
Ibn Maja*	5	6831693*	6659226	+ 2.5 %
Zolim	3	5946659	6337045	- 6%
Jaber Ibn Hayan	3	5676846	4784654	+ 18.5%
Al-Qudas	11	11368697	11525105	- 2%
Saud Bin Mosib	2	8628499	9045024	- 5%
Al-Siddiq	2	9497660	11500841	- 18%
Al-Farouq	11	12347019	14765769	- 17%
Tahfiz Al-Quran	3	7172692	6234213	+ 15%
Jafa	5	5232019	5702058	- 9%

* Concrete Area = 3453 m^2
Quantity of excavated materials = 1750 m^3
Foot type = 1
Quantity of backfilling materials = 1260m^3
No. of technical laboratories = 2

Predicted Cost = - 1257537.45 + 1721.36 (3453)
+ 55.025 (1520) + 1261193.45 (1)
+ 71.07 (1260) + 355491.7 (2)
= 6831693 (SR)

4.3.1.5 Predictability of the Model

In order to compare the predictability of the developed model against the estimated cost obtained by using the current estimating system of the Ministry of Education, a sample of schools was randomly selected. As can be seen in Table 8, the estimated cost using the developed model was greatly improved. This is thought to be due to the involvement of several influencing variables in the statistical model rather than only one variable (concrete area) as in the current estimating system used by the Ministry of Education. Another reason that might be behind the improvement obtained in the estimated cost is the great importance given to the potential cost factors in any construction project, such as the site conditions, or to the special factors which distinguish school projects from any other projects, such as its level (i.e. elementary, intermediate or secondary) and the available number of technical laboratories.

Table 8: Comparison Between the Predicted Cost Using the Developed Regression General Model and the Cost Estimate Prepared by the Ministry of Education

Project Name	Design Model	Bid Price (1)	Predicted Cost Using the Regression Model (2)	Estimated Cost Using the Ministry's System(3)	Percentage Difference Between 1 & 2	Percentage Difference Between 1 & 3
IBN MAJA	5	6659226	6831693	6042750	+2.5%	-10%
ZOLIM	3	6337045	5946659	5392400	-6%	-15%
JABER IBN HAYAN	3	4784654	5676846	5392400	+18.5	+13%
AL-QUDAS	11	11525105	11368697	9520000	-2%	-18%
SAUD BIN MOSIB	2	9045024	8628499	6510000	-5%	-28%
AL-SIDDIQ	2	11500841	9497660	6696000	-18%	-41%
AL-FAROUQ	11	14765769	12347019	9520000	-17%	-35%
THAFIZ AL-QURAN	3	6234213	7172692	5474000	+15%	-12%
JAFA	5	5702058	5232019	4438000	-9%	-22%

4.3.1.6 Confidence Level of the Model

Confidence level can be interpreted as the probability of the estimate being correct. As being known, a good estimate implies a small length of confidence interval at high level of confidence. Thus, to measure the interval at which our estimate is correct, the upper and lower confidence limit were calculated at 95% confidence level. Moreover, the percentage of the upper and lower limits of the predicted cost with respect to its mean were calculated to have an estimate of how the predicted cost may vary within the above specified level of confidence. In table 9, the predicted costs and their statistical analysis as described above are presented,

where

Predicted cost	=	The predicted cost of the project using the developed statistical model
STD.ERR.Predict	=	Standard error of estimate for the predicted cost
Lower 95% mean	=	The lower limit of the predicted cost at 95% confidence level
Upper 95% mean	=	The upper limit of the predicted cost at 95% confidence level

% of lowest pred. cost =	The percentage difference of the lower confidence limit with respect to the predicted cost at 95% confidence level.
% of upper pred. cost =	The percentage difference of the upper confidence limit with respect to the predicted cost at 95% confidence level.

By referring to table (9), it can be noticed that at 95% confidence level the predicted costs may vary with ± 12% only from their mean value. This low confidence interval is considered as an indication for the goodness and reliability of the estimated cost obtained from the developed statistical model.

TABLE 9 : CONFIDENCE LEVEL FOR THE GENERAL MODEL

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
1	5706523	109198.8	5490906	5922139	0.962215696	1.037784129
2	7685788	245735.5	7200575	8171001	0.936868803	1.063131197
3	4490999	123756.3	4246638	4735361	0.945588721	1.054411502
4	4770890	124769.7	4524528	5017252	0.948361417	1.051638583
5	6416515	176459.1	6068090	6764939	0.945698717	1.054301128
6	7167438	187954.0	6796316	7538559	0.948221108	1.051778753
7	6886191	105556.3	6677766	7094615	0.969732905	1.03026695
8	4408815	124261.9	4163456	4654175	0.944348084	1.055652142
9	6231606	105756.7	6022786	6440426	0.966490179	1.033509821
10	6530360	108892.3	6315949	6745972	0.967167047	1.033016863
11	12323307	179129.2	11969610	12677004	0.971298532	1.028701468
12	11881696	194199.8	11498241	12265150	0.96772725	1.032272665
13	11338668	207907.6	10928147	11749189	0.963794601	1.036205399
14	10844538	213662.4	10422654	11266421	0.961097098	1.03890281
15	7470782	184848.0	7105793	7835771	0.951144472	1.048855528
16	5885607	146488.3	5536361	6174853	0.940661006	1.049144634
17	6631858	104091.0	6426326	6837389	0.969008383	1.030991466
18	11527250	190479.2	11151142	11903358	0.96737227	1.03262773
19	5598645	140001.0	5322208	5875082	0.950624303	1.049375697
20	6468723	117662.0	6236396	6701051	0.964084565	1.03591559

Table 9: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
21	8417651	433691.3	7561313	9273990	0.898268769	1.10173135
22	7292070	176114.0	6944327	7639813	0.952312169	1.047687831
23	7463215	182542.7	7102778	7823651	0.951704862	1.048295004
24	4433832	125113.5	4186791	4680873	0.944282733	1.055717267
25	3954743	128159.2	3701688	4207798	0.936012277	1.063987723
26	6638350	109503.3	6422132	6854568	0.967428954	1.032571046
27	12444823	159857.2	12129180	12760467	0.974636602	1.025363478
28	12392322	175593.1	12045607	12739036	0.972021789	1.02797813
29	12547396	229244.4	12094745	13000047	0.963924706	1.036075294
30	12404003	155187.0	12097581	12710425	0.975296523	1.024703477
31	6579572	119480.5	6343654	6815491	0.964143868	1.035856284
32	11194544	231237.3	10737958	11651130	0.959213524	1.040786476
33	5380727	124706.2	5134490	5626964	0.954237225	1.045762775
34	5639353	129940.9	5382780	5895926	0.954503114	1.045496886
35	7144690	183318.1	6782722	7506658	0.94933748	1.05066252
36	7181991	177622.5	6831269	7532713	0.951166466	1.048833534
37	6542095	112097.1	6320755	6763434	0.966166801	1.033833046
38	7128222	178819.8	6775136	7481308	0.95046647	1.04953353
39	7241949	176065.3	6894302	7589596	0.951995381	1.048004619
40	7807584	137733.0	7535626	8079543	0.965167458	1.03483267
41	5720833	151657.3	5421381	6020286	0.947655875	1.0523443
42	4735673	128709.3	4481532	4989814	0.946334766	1.053665234

Table 9: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
43	7344244	192972.5	6963213	7725275	0.948118418	1.051881582
44	7334580	174635.9	6989755	7679404	0.952986401	1.047013462
45	6543578	110602.3	6325190	6761967	0.966625598	1.033374554
46	6885192	221881.3	6447080	7323305	0.936368949	1.063631196
47	12605633	220268.4	12170705	13040560	0.96549733	1.034502591
48	5775189	105297.8	5567274	5983103	0.963998581	1.036001246
49	6579699	114060.7	6354483	6804916	0.965771079	1.034229073
50	7827131	378678.4	7079417	8574844	0.904471511	1.095528362
51	13508590	255233.9	13004622	14012558	0.962692775	1.037307225
52	6702467	127277.7	6451152	6953781	0.962504105	1.037495746
53	12486746	230464.0	12031687	12941805	0.963556638	1.036443362
54	5085012	148361.0	4792069	5377956	0.942390893	1.057609304
55	4581267	118201.8	4347874	4814661	0.949054923	1.050945295
56	6416514	120902.7	6177788	6655241	0.962795063	1.037205093
57	7715874	197090.5	7326712	8105036	0.949563458	1.050436542
58	8009885	171776.9	7670705	8349064	0.957654823	1.042345052
59	6127284	109037.4	5911986	6342582	0.964862409	1.035137591
60	13264485	187825.0	12893618	136335352	0.972040603	10.2782243
61	13493745	256570.3	12987138	14000352	0.96245616	1.03754384
62	12944717	273263.8	12405148	13484285	0.958317436	1.041682487
63	8295919	164464.4	7971178	8620659	0.960855331	1.039144548
64	11993469	179981.2	11638090	12348848	0.970368957	1.029631043

Table 9: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
65	4784723	126985.6	4533985	5035460	0.94759613	1.052403661
66	7611109	199842.4	7216513	8005704	0.948155256	1.051844613
67	7332515	174611.6	6987738	7677291	0.952979708	1.047020156
68	4629761	117333.2	4398082	4861439	0.949958756	1.050041028
69	12697088	303058.7	12098689	13295488	0.952871162	1.047128916
70	9105560	295762.1	8521568	9689553	0.935864241	1.064135869
71	5643081	224875.0	5199057	6087104	0.921315324	1.078684499
72	11414469	239504.9	10941558	11887379	0.958569163	1.041430749
73	4378092	126805.1	4127711	4628473	0.942810475	1.057189525
74	11561179	192052.2	11181965	11940393	0.967199366	1.032800634
75	10729405	200929.3	10332663	11126147	0.963022926	1.036977074
76	12540714	280043.4	11987758	13093669	0.955907136	1.044092785
77	12271476	159690.9	11956161	12586791	0.974305047	1.025694953
78	12631684	155174.6	12325286	12938082	0.975743693	1.024256307
79	5992006	120385.9	5754300	6229712	0.960329479	1.039670521
80	6504263	115885.4	6275443	6733083	0.964819996	1.035180004
81	5987008	121409.8	5747280	6226736	0.95995863	1.04004137
82	6542770	127029.9	6291945	6793595	0.961663791	1.038336209
83	5983843	120959.6	5745004	6222682	0.960086018	1.039913982
84	5411561	121409.5	5171834	5651288	0.955700952	1.044299048
85	7049446	116126.5	6820151	7278742	0.967473331	1.032526811
86	4532518	124062.7	4287552	4777484	0.945953662	1.054046338

Table 9: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
87	16437371	739962.1	14976290	17898452	0.911112245	1.088887755
88	13195443	254027.7	12693857	13697030	0.961987938	1.038012138
89	12700671	210993.8	12284057	13117286	0.967197481	1.032802598
90	13687065	217004.6	13258582	14115548	0.968694311	1.031305689
91	13092732	213806.9	12670563	13514902	0.967755469	1.032244607
92	13716681	423150.4	12881156	14552206	0.939086941	1.060913059
93	13105343	214321.7	12682157	13528528	0.967708896	1.032291028
94	13643509	246318.6	13157145	14129874	0.964351986	1.035648087
95	12509443	226854.1	12061512	12957374	0.96419257	1.03580743
96	5503051	124495.8	5257230	5748872	0.955330052	1.044669948
97	6452307	118892.7	6217549	6687064	0.963616424	1.036383421
98	5433577	123541.4	5189640	5677514	0.955105633	1.044894367
99	4883001	130993.3	4624350	5141652	0.94703032	1.05296968
100	3891463	130706.4	3633379	4149548	0.933679441	1.066320816
101	8161144	172961.4	7819626	8502662	0.958153171	1.041846829
102	6022870	120188.1	5785555	6260186	0.960597688	1.039402478
103	6657881	103091.7	6454323	6861439	0.969426008	1.030573992
104	4405721	125365.6	4158182	4653260	0.943814191	1.056185809
105	5507618	135735.9	5239603	5775633	0.951337402	1.048662598
106	5358252	122916.4	5115549	5600954	0.954704818	1.045294995
107	6200417	123166.0	5957222	6443613	0.960777638	1.039222523
108	6730898	125488.5	6483117	6978680	0.963187527	1.036812621

Table 9: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
109	4421769	122938.4	4179023	4664515	0.945102062	1.054897938
110	6456224	114760.4	6229625	6682822	0.96490224	1.035097605
111	7497332	108668.6	7282762	7711902	0.971380486	1.028619514
112	5997509	119710.7	5761136	6233882	0.960588138	1.039411862
113	4755071	130304.2	4497781	5012362	0.945891449	1.054108761
114	4939189	142102.4	4658603	5219775	0.943191888	1.056808112
115	4881684	139428.1	4606379	5156990	0.943604502	1.056395703
116	5537318	135860.2	5269057	5805578	0.951553983	1.048445836
117	5374997	130632.8	5117058	5632936	0.952011322	1.047988678
118	4215894	142577.4	3934370	4497418	0.933223179	1.066776821
119	7362241	200384.1	6966576	7757906	0.946257532	1.053742468
120	7197585	180458.6	6841264	7553907	0.950494367	1.049505772
121	6595061	112371.3	6373180	6816942	0.96635649	1.03364351
122	6791549	110761.4	6572846	7010251	0.967797773	1.032202079
123	6412430	118699.7	6178053	6646807	0.963449582	1.036550418
124	4451922	123852.4	4207372	4696473	0.945068669	1.054931555
125	6772273	191646.2	6393861	7150685	0.944123339	1.055876661
126	8425344	373850.8	7687163	9163526	0.912385655	1.087614464
127	7560347	180575.9	7203794	7916901	0.952839069	1.047161063
128	5868374	379690.0	5118662	6618085	0.872245361	1.127754468
129	7836818	155887.3	7529013	8144623	0.960723217	1.039276783
130	7172475	177538.7	6821919	7523032	0.951124821	1.048875318

Table 9: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
131	4795878	120507.8	4557932	5033825	0.95038531	1.049614898
132	10089030	166766.4	9759744	10418316	0.967361976	1.032638024
133	4565335	123680.0	4321125	4809546	0.946507759	1.05349246
134	12624349	154466.2	12319350	12929348	0.975840418	1.024159582
135	10067247	165847.9	9739775	10394719	0.967471544	1.032528456
136	11799716	353369.4	11107976	12497456	0.941376555	1.059131932
137	5419148	125165.6	5172004	5666291	0.954394307	1.045605508
138	7264805	175591.4	6918094	7611517	0.95227525	1.047724887
139	6749355	131062.4	6490568	7008142	0.961657521	1.038342479
140	6750728	145732.4	6462974	7038481	0.957374375	1.042625477
141	5518257	135215.4	5251269	5785244	0.951617331	1.048382487
142	7045148	118816.7	6810541	7279756	0.966699493	1.033300649
143	5670861	111817.2	5450074	5891648	0.961066406	1.038933594
144	5789515	112515.6	5567349	6011681	0.961626147	1.038373853
145	4885536	138957.4	4611160	5159912	0.94383912	1.05616088
146	6417130	118252.7	6183636	6650624	0.963613952	1.036386048
147	6453245	116846.5	6222528	6683963	0.964247909	1.035752246
148	7294207	111580.4	7073888	7514526	0.969795346	1.030204654
149	6451640	116508.8	6221589	6681691	0.964342245	1.035657755
150	6599474	117758.2	6366957	6831992	0.964767344	1.035232808
151	7202147	115465.3	6974157	7430137	0.968344162	1.031655838
152	7234360	109114.9	7018908	7449811	0.970218236	1.029781625

Table 9: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
153	6463496	118638.8	6229240	6697752	0.963757075	1.036242925
154	8480994	377183.3	7736232	9225755	0.912184586	1.087815296
155	4500904	122495.8	4259032	4742776	0.946261462	1.053738538
156	6154757	113933.5	5929791	6379723	0.963448435	1.036551565
157	5749853	204886.0	5345299	6154408	0.929640984	1.07035919
160	6711999	115781.8	6483384	6940614	0.965939357	1.034060643
161	3949016	132190.5	3688002	4210031	0.933904041	1.066096212
162	6510514	114805.8	6283826	6737202	0.965181244	1.034818756
163	4468546	124431.7	4222852	4714241	0.945017014	1.054983209
164	3937322	131092.0	3678476	4196168	0.934258361	1.065741639
165	7160330	187961.4	6789193	7531466	0.948167612	1.051832248
166	6619247	104727.9	6412458	6826036	0.968759437	1.031240563
167	6160952	298978.9	5570608	6751296	0.904179744	1.095820256
168	14977766	519069.2	13952846	16002686	0.931570569	1.068429431
169	14098651	331280.4	13444526	14752775	0.953603717	1.046396212
170	11330984	204668.1	10926860	11735108	0.964334607	1.035665393

4.3.1.7 Fitting of Regression Line

As is known, the regression function represented by its plotted line provides a good fit to the sample data if it results in small residuals. So, to measure the degree of fitting of the developed model to the sample data, the residual, listed in Table 6 were carefully studied and analyzed. Moreover, the regression line was also plotted as seen in Figure (11).It was found that the bid price data represented by their plotted points were distributed on both sides of the fitted line. Moreover, the dispersion was found to be small and in an average range of 9% of the bid price. Hence it can be concluded that the developed regression function provides a good fit to the sample data.

Since each design model has its own characteristics and building requirements which might introduce a sort of variation in their cost, it was decided to develop a statistical model for each different design model separately. The development of such models will reduce the large variations in the data points and is thought to end up with more accurate results. In the following sections, the regression models for school design models number 2,3,5 and 11 will be developed and presented.

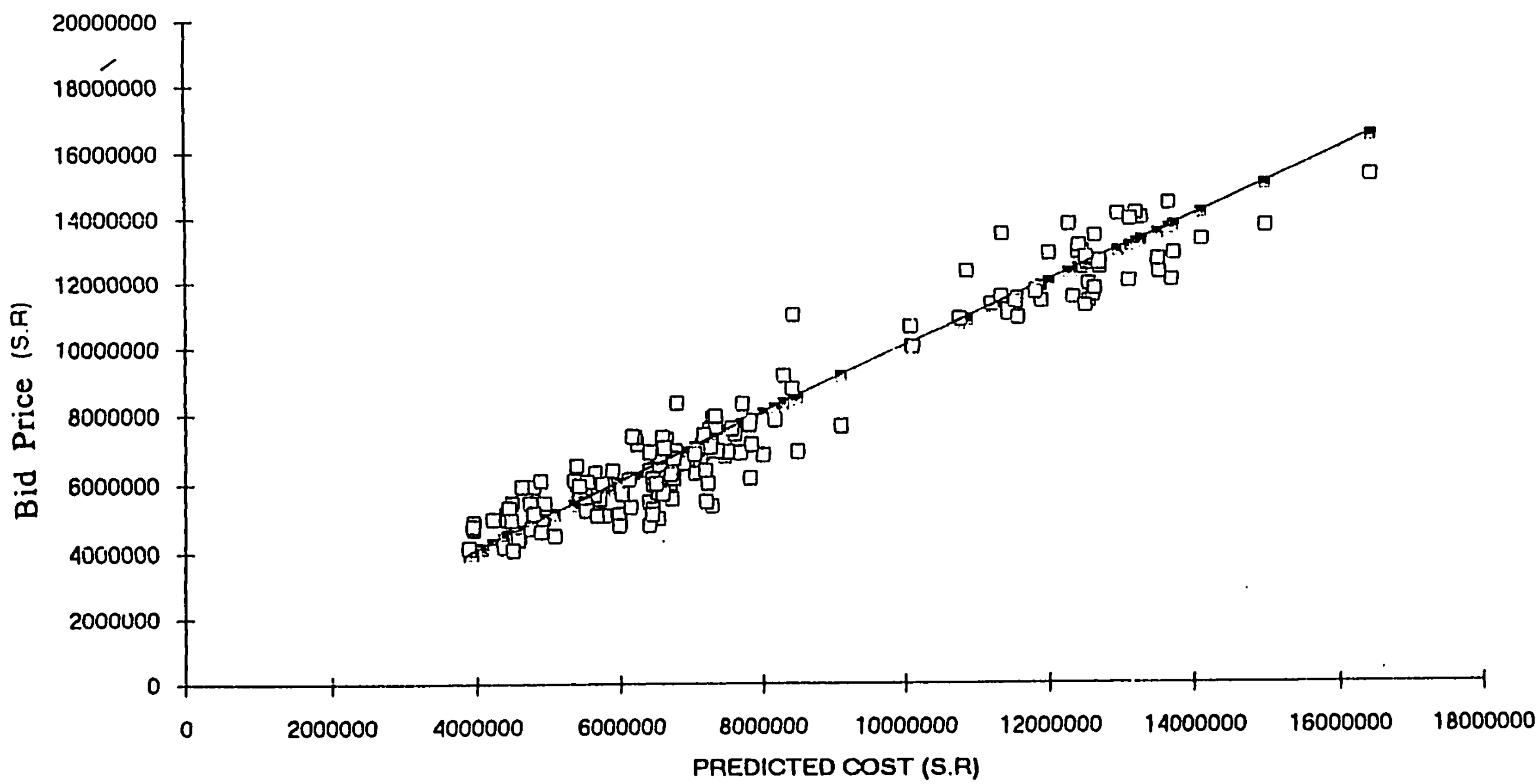

Figure 11: Regression Line for General Model

4.3.2 School Design Model # 2

4.3.2.1 Determining the probability distribution parameters

Since it was decided to develop a statistical cost estimate model for each school design model separately, a sample of bid prices for 77 boys schools submitted to the Ministry of Education in 1994 was used to determine the best linear regression function for school design model number two. The general form for the potential regression equation is as follows:

Predicted Cost $= \beta_0 + \beta_1 (X_1) + \beta_2 (X_2) + \beta_3 (X_3) +\ldots\ldots \beta_n = (X_n)$
where X_1, X_2, X_3 **and** X_n are the independent variables used as concrete area, number of classrooms, bearing capacity of soil, etc.

The following assumptions are made:

1. The construction cost is a linear function of the independent variables.

2. No relation exists between the independent variables.

The validation of the first assumption was checked by analyzing and studying the residual plot, while the validation of the second assumption was performed by studying the correlation matrix of all the independent variables shown in Appendix B. The linear least square regression was performed using the step wise procedure available in the **SAS** package to calculate the regression coefficients, β_k, and the variance of estimates, σE^2, of the construction cost regression equations.

After examining the correlation matrix, the variables found with high correlation factors (IRI > 0.5) are as follows:

Execution time	-	Number of class rooms
Bearing capacity	-	Footing type
Soil improvement	-	Footing type
Length of retaining wall	-	Quantity of backfilling material
Length of retaining wall	-	Quantity of excavated material
Level of school	-	Number of technical laboratories
Concrete area	-	Number of class rooms
Quantity of excavated material	-	Footing type

* Number of class rooms, footing type, length of retaining walls and level of school were dropped.

The different regression models were developed to investigate which set of non-related independent variables produces the needed model with higher R^2. This analysis indicated that the following variables produce the highest R^2:

1. Soil improvement
2. Number of technical laboratories
3. Quantity of excavated material
4. Quantity of backfilling material
5. length of fence
6. Concrete area
7. Soil bearing capacity
8. Execution time

Therefore, the potential form for the general potential regression equation for the construction cost of school project is as follows:

CC = $\beta_0 + \beta_1$ **(soil improvement) +** β_2 **(Number of technical laboratories) +** β_3 **(quantity of excavation materials) +** β_4 **(quantity of backfilling materials) +** β_5 **(length of fence) +** β_6 **(concrete area), +** β_7 **(soil bearing capacity) +** β_8 **(execution time).**

where:

CC = school construction cost

β_0 = regression constant

β_k = factor weights ,K = 1,.....8

The liner least regression was performed using the step wise procedure available in the **SAS** package to calculate the regression coefficient, β_k, and variance of estimates, σ_E^2 of the construction cost regression equation (Table 10). The regression analysis correlates the 77 schools' construction cost to the eight potential factors. The R^2 value for the regression is 59.66 percent. The eight factors explain 59.66 percent of the variation in the schools construction cost.

TABLE 10 : REGRESSION COEFFICIENTS AND THEIR SIGNIFICANCE LEVEL FOR SCHOOL DESIGN MODEL # 2

Regression Equation :-

$CC = \beta_0 + \beta_1$ (Soil improvement) + β_2 (Lab. No) + β_3 (quantity of excavation materials) + β_4 (quantity of backfilling materials) + β_5 (length of fence) + β_6 (concrete area) + β_7 (soil bearing capacity) + β_8 (execution time)

Coefficient of Multiple Determination $R^2 = 0.5966$

Multiple Correlation Coefficient R = 0.7723

Source Of Variation	Degree of Freedom	Sum Of Squares	Mean Squares	F Ratio	Significant At
Regression	8	62032717967179	7754089745897	12.57	99.99%
Residual	68	41932815033558	616659044611	-	-

Variable	Regression Coefficient	Standard Error of Estimate σ_E	Variance of Estimate σ_E2	T-Ratio	Significant At
Constant	-65703902	74994873	5.549486178E15	0.87	61%
Soil Improvement	797026.83	331496.93	1.098902146E11	2.40	98.11%
Bearing Capacity	-249260.4	228378.32	5.215665705E10	1.09	72.11%
Execution Time	4758419.9	5314681.76	2.824584221E13	.89	62.62%
Lab No.	518590.37	311445.01	550.550	1.66	89.95%
Quantity of Beackfilling Material	57.70	23.464	9.699799425E10	2.45	98.35%
Quantity of Excavated Matrial	76.19	28.130	791.2969	2.71	99.15%
Length of Fence	5572.19	1455.11	2117345.112	3.82	99.97%
Concrete Area	-11963.94	14320.15	205066696	0.83	59.36%

TABLE 11 : REGRESSION COEFFICIENTS AND THEIR SIGNIFICANCE LEVEL FOR SCHOOL DESIGN MODEL # 2

Regression Equation :-

$CC = \beta_0 + \beta_1$ (Soil improvement) + β_2 (Lab. No) + β_3 (fill) + β_4 (excavation) + β_5 (fence) + β_6 (concrete area)

Coefficient of Multiple Determination $R^2 = 0.5864$

Multiple Correlation Coefficient $R = 0.7657$

Source Of Variation	Degree of Freedom	Sum Of Squares	Mean Squares	F Ratio	Significant At
Regression	6	60973674597859	10162279099643	16.55	99.99%
Residual	70	42991858402878	614169405755	-	-

Variable	Regression Coefficient	Standard Error of Estimate σ_E	Variance of Estimate σ_E^2	T-Ratio	Significant At
Constant	1155911.57	727014.64	5.285502868E11	1.59	88.36%
Soil Improvement	1006658.61	288489.54	8.322621469E10	3.48	99.92%
Lab. No.	235291.46	87119.3	7589772432	2.7	99.13%
Quaantity of Backfilling Material	55.33	23.346	545.03	2.37	97.95%
Quantity of Excavated Material	74.36	27.94	780.643	2.66	99.04%
Length of Fence	5231.75	1428.84	2041583.74	3.66	99.95%
Concrete Area	849.95	171.16	29295.74	4.96	99.99%

Since it is risky to consider some of the variables which have low significance level such as the bearing capacity, execution time and the concrete area, it was decided to go back among the different statistical models developed using the step wise procedure until the significance level of the selected parameters became higher than 95%. The selected statistical model is presented in Table 11.

Since the confidence level for the selected model represented by its significance level is considered very high (99.99%), it can be concluded that the model is statistically significant. Regarding the parameters themselves, it can be seen that their confidence level is also very high (97.95% - 99.99%) and we can conclude that all the above parameters will be considered in the selected linear statistical model.

Thus, the linear statistical model for school design model number 2 built by the Ministry of Education in Saudi Arabia is as follows:

Predicted cost = 1155911.57+ 1006658.61 (soil - imp) + 235291.46 (Lab No.) + 55.33 (fill) + 74.36 (Excavation) + 5231.75 (fence) + 849.95 (Concrete area)

where

Soil - Imp	=	Soil improvement
Lab No.	=	Number of technical laboratories available in school.
Fill	=	Quantity of backfilling material (m^3)
Excav	=	Quantity of excavated material (m^3)
Fence	=	Length of fence (m)
Conc-area	=	Concrete area (m^2)

As can be seen from the above statistical model, six different independent variables significantly contribute to the cost of boys schools built as per school design model number 2. Those factors are classified into two categories. First, the substructure category including soil improvement and quantity of backfilling and excavation materials. Second, the superstructure category including the concrete area, length of fence and the number of technical laboratories available in the school. Having both the categories included in the developed model indicates that they are almost equally important in predicting the cost of boys schools built as per school design model number 2.

In addition to the other independent variables that were previously discussed in the general statistical model and found to have a great impact on the cost of school projects such as the quantity of excavated and backfilling materials, concrete area and number of technical laboratories available in school, two more new independent variables come into the picture for school design model number 2. These two variables are soil improvement and the length of fence.

Contrary to what was found in the general model, the effect of soil improvement had a very important impact on the total construction cost of schools built according to school design model number 2. It was found from the collected data that about 10% of the schools had soil improvement for their site. Moreover, those schools were built in the Al-Qatif and Al-Hassa areas where the problems of swelling and planetary soils are prevalent. The cost of improving swelling soils can be very expensive since it can involve very advanced and complicated techniques. Accordingly, the effect of improving the soil in those areas was expensive and had the greatest impact on the total construction cost of boys schools built as per design model number 2. This great impact was found to be an additional amount of 17.5% of the total construction cost of the project if the soil had to be improved.

The other independent variable which was found to have an impact on the construction cost of school projects built as per school design model number 2 is the length of fence. Normally the cost of fence is dependent on its height , length and the type of footing used (i.e. isolated or strip footing). Since no detailed information was available for the researcher on some of these factors, the considerable effect of fence on the total cost of school design model number 2 is thought to be due to either its length or its type of footing (i.e. the use of strip footing rather than the quoted one) . The effect of this factor on the total construction cost of the school project was found to be 0.9% for every additional 10 (m).

Table 12: The Average Effect of Increase in one of the Parameters on the Overall Construction Cost of School Project for school design model number 2.

Parameter	Amount of Increase	Percent of Increase
Soil improvement	done	17.5%
Number of technical laboratories	2	8%
Quantity of filling material	100m3	0.09%
Quantity of excavated material	$100m^3$	0.125%
Length of fence	10m	0.9%
Concrete area	$100m^2$	1.8%

4.3.2.2 Testing of Assumption

When a regression model is applied in practice, one cannot usually be sure in advance that the model implied is appropriate for the situation in hand. Consequently, the model considered needs to be checked for suitability and its agreement with the earlier assumptions made to develop it. As stated earlier, the following assumption were made:

i The regression function is linear

ii The distribution is normal

iii There is a constant variance over all schools construction cost.

In the following paragraphs, the above mentioned assumptions will be verified and checked.

4.3.2.2.1 Linearity Assumption

In order to test the developed model for its aptness and linearity, the residual plot was analyzed and studied. This graph was obtained by plotting the residuals against the fitted values (Figure 12). If the specified regression model is linear, the residuals will tend to scatter at random around the zero line when plotted against the fitted values. As can be seen in Figure 12 there is no pattern of systematic departure and all points are scattered normally around the zero line. This shows that the selected statistical model for all design models built by the Ministry of Education is linear and apt.

4.3.2.2.2 Normality and Constant Variance Assumptions

In order to check the normality assumption, and to be sure that a constant variance exists over all schools construction cost, the lack of normality test was performed. In this test, the standardized residuals were calculated. If the distribution was normal with the constant variance and the sample size is considerably large (greater than 30 observations), the standardized residuals will tend to follow the standard normal distribution (figure 10). Hence, about half of the standardized residuals should be positive and half negative, About 68% of the standardized residual distribution should fall between -1 and +1 and at least 95% of the area under the normal distribution should be contained within t (0.025: n-2) and t (0.975: n-2) (i.e. $\mu \pm 2\sigma$) (Netir, Warerman and Whitemore, 1978).

In order to check if the developed model complies with the above mentioned criteria of normal distribution, the standardized residuals were calculated and listed in table 13.

Based on the analysis that was conducted on the results listed in table 13, the following was found

$$t \quad (0.025 : 75) = -1.99$$

$$t \quad (0.975 : 75) = +1.99$$

$$-1 < 76.6\ \% \text{ of the distribution} < +1$$

57% of the distribution are negative

$$-1.99 < 97.4\% \text{ of the distribution} < 1.99$$

Since the above analysis complies with the different test criteria, then it can be concluded that this analysis provides no evidence of any departure form normality and the variance is constant over all school construction cost.

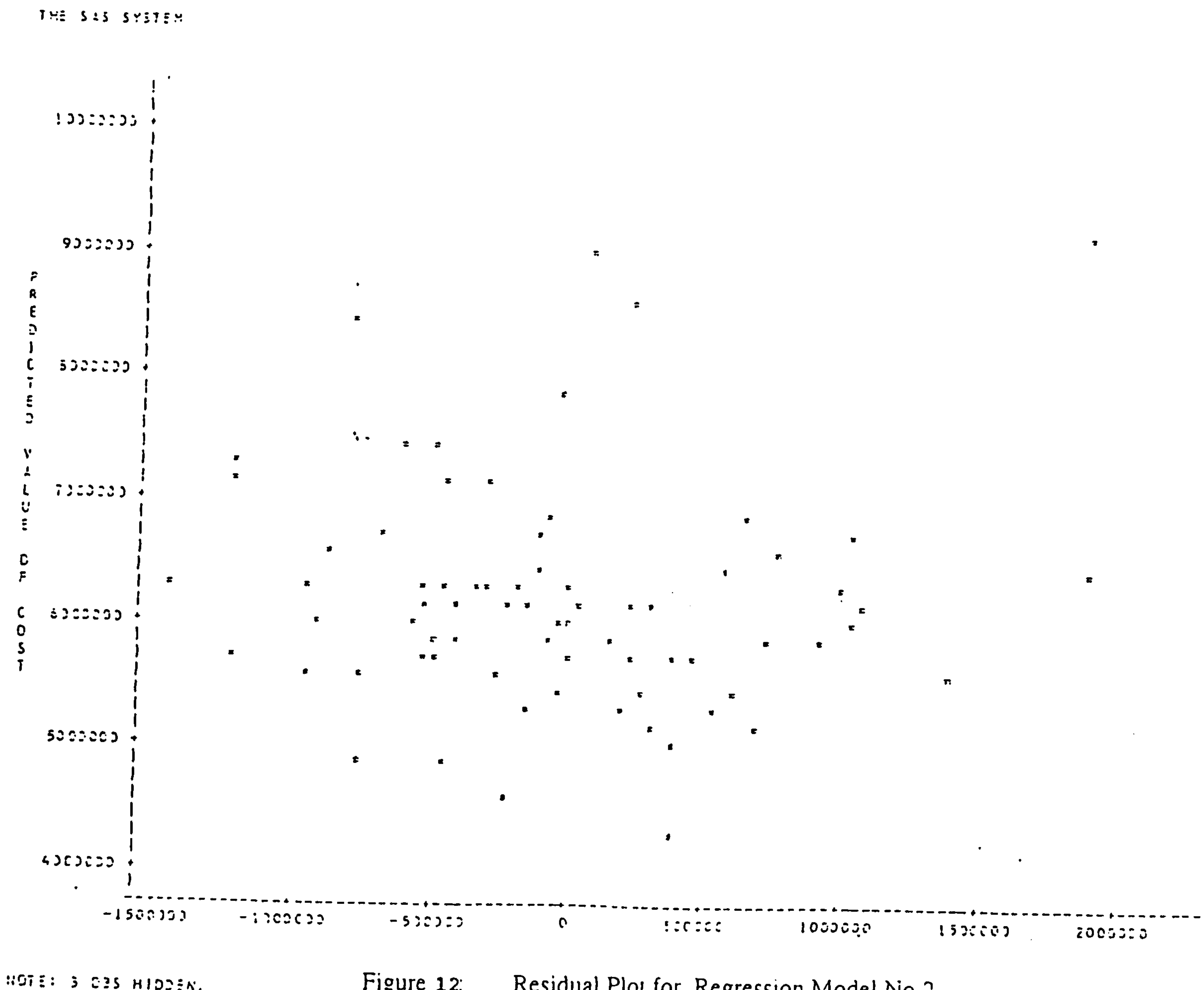

Figure 12: Residual Plot for Regression Model No.2

TABLE 13 : RESIDUAL TABLE FOR REGRESSION MODEL # 2

NO	MODEL	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
1	2	6904694	7206007	-301313	-0.384472474
2	2	5436815	5041028	395787	0.505020384
3	2	5733160	5455296	277864	0.354551776
4	2	6392235	6096818	295417	0.376949235
5	2	5874050	6106270	-232220	-0.296310474
6	2	5868298	5344967	523331	0.667765294
7	2	5329678	5585791	-256113	-0.326797711
8	2	8742473	8527482	214991	0.274326437
9	2	7268715	6203238	1065477	1.359538347
10	2	5833170	6309902	-476732	-0.608305421
11	2	5950785	5704472	246313	0.314293006
12	2	6337282	6446335	-109053	-0.139150573
13	2	6819579	5902036	917543	1.170775994
14	2	7696680	6664160	1032520	1.317485533
15	2	5955973	5790429	165544	0.211232543
16	2	5428406	5219117	209289	0.26705074
17	2	6395654	6183978	211676	0.270096529
18	2	6594874	6697749	-102875	-0.131267505
19	2	5862659	7077224	-1214565	-1.549773193
20	2	6134082	7348973	-1214891	-1.550189166
21	2	5985308	6696767	-711459	-0.907814803
22	2	4453455	5693015	-1239560	-1.581666571

Table 13: Continued

NO	MODEL	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
23	2	4343296	4586412	-243116	-0.310213665
24	2	5464115	5144876	319239	0.407345876
25	2	6832657	6897261	-64604	-0.082434079
26	2	9125940	9061336	64604	0.082434079
27	2	5895825	5192460	703365	0.897486937
28	2	7655616	8449019	-793403	-1.012374554
29	2	5660981	6550250	-889269	-1.134698643
30	2	5758779	5831961	-73182	-0.093379524
31	2	5718324	5730771	-12447	-0.015882252
32	2	5723083	6275135	-552052	-0.704413013
33	2	5885340	6247410	-362070	-0.461997818
34	2	5189223	5704934	-515711	-0.658042248
35	2	5686894	6243573	-556679	-0.71031702
36	2	4813824	6260131	-1446307	-1.845473743
37	2	7821456	7849958	-28502	-0.036368276
38	2	6181035	5736501	444534	0.567221084
39	2	5431683	5915380	-483697	-0.617192694
40	2	6095409	5717169	378240	0.482630582
41	2	5538262	6080536	-542274	-0.691936379
42	2	6130537	6324489	-193952	-0.247480876
43	2	4673077	4277387	395690	0.504896613
44	2	5437666	5471038	-33372	-0.042582349
45	2	6083932	5471561	612371	0.781379473

Table 13: Continued

NO	MODEL	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
46	2	6045409	5658019	387390	0.494305893
47	2	6529180	5796725	732455	0.934605496
48	2	7337905	6339934	997971	1.273401344
49	2	8335448	6435741	1899707	2.424007758
50	2	6908921	5541060	1367861	1.745377406
51	2	6957923	6386980	570943	0.728517746
52	2	10990783	9118063	1872720	2.389572607
53	2	7577367	6918490	658877	0.840720679
54	2	8850205	6249920	2600285	3.31793851
55	2	5118840	5664994	-546154	-0.696887222
56	2	4350750	4820565	-469815	-0.599479396
57	2	5932772	5683536	249236	0.318022725
58	2	6156745	6108130	48615	0.06203227
59	2	6749661	7198954	-449293	-0.573293522
60	2	5203988	5347346	-143358	-0.182923421
61	2	6873643	7480650	-607007	-0.774535061
62	2	4587976	5536700	-948724	-1.210562648
63	2	4761702	5514130	-752428	-0.960090851
64	2	5251840	6216423	-964583	-1.230798579
65	2	5342725	5930842	-588117	-0.750431604
66	2	5103970	6034389	-930419	-1.187205645
67	2	5695288	6125461	-430173	-0.548896587
68	2	5463114	5890264	-427150	-0.545039269

Table 13: Continued

NO	MODEL	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
69	2	5991870	6284425	-292555	-0.37329735
70	2	5853221	5961545	-108324	-0.138220376
71	2	6935347	7441045	-505698	-0.645265757
72	2	4036257	4802506	-766249	-0.977726313
73	2	5986049	6142855	-156806	-0.200082939
74	2	6259610	6255833	3777	0.004819415
75	2	6007055	6051994	-44939	-0.057341729
76	2	7036454	5992580	1043874	1.331973128
77	2	7386606	6612903	773703	0.987237545

4.3.2.3 *Validity of the Model:*

In order to test or to verify the validity of the model, six schools were selected randomly, excluded from the data base and withheld for checking the predictability of the developed model. The reliability of the models predictability was examined by comparing the bid prices with the model estimates (Table 14). As can be seen from Table 14 that the developed model tends to under estimate the project cost. On the other hand, the model possesses a balanced behavior when it over or under estimates the project costs (i.e. the percent difference between the bid price and predicted cost reached very close values with positive and negative signs). This balanced action adds more credits to the reliability and predictability of the developed statistical model.

Moreover, by considering the fact that the model is primarily designed for the purpose of preliminary estimating and the variation in predicted and bid price costs falls within the acceptable limits (-30%, + 30%), then it can be concluded that the developed model is the most representative statistical model for school design model # 2.

Table 14: A Comparison between the Bid Price and the Predicted Cost for School Design model #2.

Project Name	Predicted Cost	Bid Price	Percent Difference
Sohib Al-Romi*	6243989.67*	7322704	-15%
Otqba Bin Massod	5078633.9	4165781	+21%
Saad Bin Mosib	6838718.74	9045024	-24%
Al-Isamla	5380514.54	5901360	-9%
Moditha	6512107	5958347	+8%
Al-Sidiq	9060340.66	11500841	-21%

* Soil Improvement = 0
Lab No. = 0
Quantity of backfilling materials = 6276 m^3
Quantity of excavated materials = 0m^3
Length of fence = 310m
Concrete area = 3670m^2

Predicted cost = 1155911.57 + 1006658.61(0)
+ 1235291.46 (0) + 55.33 (6276)
+ 74.36 (0) + 5231 (310)
+ 849.95 (3670) = 6243989.67SR

4.3.2.4 Predictability of the Model

In order to compare the predictability of the developed model against the estimated cost obtained by using the current estimating system of the Ministry of Education, a sample of schools was randomly selected. As can be seen in Table 15, the estimated cost using the developed model was greatly improved. This is thought to be due to the involvement of several influencing variables in the statistical model rather than only one variable (concrete area) as in the current estimating system used by the Ministry of Education. Another reason that might be behind the improvement obtained in the estimated cost is the great importance given to the potential cost factors in any construction project, such as the site conditions, or to the special factors which distinguish school projects from any other projects, such as its level (i.e. elementary, intermediate or secondary) and the available number of technical laboratories.

Table 15: Comparison Between the Predicted Cost Using the Developed Regression Model # 2 and the Cost Estimate Prepared by the Ministry of Education

Project Name	Design Model	Bid Price (1)	Predicted Cost Using the Regression Model (2)	Estimated Cost Using the Ministry's System(3)	Percentage Difference Between 1 & 2	Percentage Difference Between 1 & 3
SOHIB AL-ROUMI	2	7322704	6243989	6239000	-15%	-15%
OTBA BIN MASSOD	2	4165781	5078633	4340100	+21%	+4%
SAAD BIN MOSIB	2	9045024	6838718	6510000	-24%	-28%
AL-ISMALA	2	5901360	5380514	4555250	-9%	-22%
MOHDITHA	2	5958347	6512107	6696000	+9%	+12%
AL-SIDIQ	2	11500841	9060340	6696000	-21%	-42%

4.3.2.5 *Confidence Level of the Model*

Confidence level can be interpreted as the probability of the estimate being correct. As being known, a good estimate implies a small length of confidence interval at high level of confidence. Thus, to measure the interval at which our estimate is correct, the upper and lower confidence limit were calculated at 95% confidence level. Moreover, the percentage of the upper and lower limits of the predicted cost with respect to its mean were calculated to have an estimate of how the predicted cost may vary within the above specified level of confidence. In table 17, the predicted costs and their statistical analysis as discussed above are presented.

By referring back to table (17), it can be noticed that the bid price may vary from the predicted value as listed in Table (16) below.

Table (16): Variation of Bid Price with Respect to the Predicted Cost school for design model number 2

Percent of the total observation points	Range of Variation at 95% confidence level
86%	±10%
12%	±15%
2%	±20%

From the above results, it can be noticed that at 95% confidence level, the deference between the predicted costs and their means for the majority of the observation points (98%) is below 15% while 2% only of the observation points reached upto 15%. This low confidence interval is considered as an indication for the goodness and reliability of the estimated cost obtained from the developed statistical model.

4.3.2.6 Fitting of Regression Line

As is known, the regression function represented by its plotted line provides a good fit to the sample data if it results in small residuals. So, to measure the degree of fitting of the developed model to the sample data, the residuals listed in Table 13 were carefully studied and analyzed. Moreover, the regression line was also plotted as seen in Figure 13. It was found that the bid price data represented by their plotted points were distributed on both sides of the fitted lines. Moreover, the dispersion was found to be small and in an average range of 8 % of the actual cost. Hence it can be concluded that the developed regression function provides a good fit to the sample data.

TABLE 17 : CONFIDENCE LEVE FOR REGRESSION MODEL # 2

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	%OF LOWEST PRED.COST	%OF HIGHEST PRED.COST
1	7206007	377440.2	6452836	7959177	0.895480118	1.104519743
2	5041028	204893.1	4632170	5449886	0.918893924	1.081106076
3	5455296	199344.3	5057511	5853081	0.927082783	1.072917217
4	6096818	194687.2	5708326	6485310	0.936279548	1.063720452
5	6106270	156250.4	5794478	6418063	0.948939041	1.051061122
6	5344967	228406.0	4889190	5800744	0.914727818	1.085272182
7	5585791	176689.5	5233212	5938369	0.936879307	1.063120514
8	8527482	523054.0	7483744	9571220	0.877603025	1.122396975
9	6203238	169691.6	5864624	6541853	0.945413347	1.054586814
10	6309902	186584.1	5937579	6682225	0.940993854	1.059006146
11	5704472	198883.0	5307607	6101337	0.930429144	1.069570856
12	6446335	223814.9	5999719	6892951	0.930717842	1.069282158
13	5902036	149591.8	5603530	6200542	0.949423216	1.050576784
14	6664160	288305.3	6088855	7239464	0.913671791	1.086328059
15	5790429	219919.0	5351587	6229270	0.924212524	1.075787303
16	5219117	186148.2	4847664	5590570	0.928828382	1.071171618
17	6183978	175198.3	5834376	6533581	0.943466487	1.056533675
18	6697749	268856.3	6161254	7234243	0.919899208	1.080100643
19	7077224	280401.6	6517692	7636757	0.920939057	1.079061084
20	7348973	291527.6	6767239	7930708	0.920841456	1.07915868
21	6696767	238349.3	6221148	7172386	0.928977819	1.071022181
22	5693015	215356.4	5263278	6122752	0.924515042	1.075484958

Table 17: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	%OF LOWEST PRED.COST	%OF HIGHEST PRED.COST
23	4586412	222895.9	4141630	5031194	0.903021796	1.096978204
24	5144876	228314.6	4689282	5600471	0.91144704	1.088553155
25	6897261	584334.1	5731240	8063281	0.830944342	1.169055513
26	9061336	584334.1	7895316	10227357	0.871319196	1.128680914
27	5192460	194058.7	4805222	5579698	0.925423017	1.074576983
28	8449019	462444.9	7526224	9371813	0.890780811	1.10921907
29	6550250	312772.1	5926123	7174377	0.904717072	1.095282928
30	5831961	167221.8	5498275	6165647	0.942783225	1.057216775
31	5730771	237979.3	5255891	6205651	0.917135059	1.082864941
32	6275135	246801.4	5782651	6767619	0.921518182	1.078481818
33	6247410	234378.8	5779714	6715105	0.925137617	1.074862223
34	5704934	171868.8	5361975	6047893	0.939883792	1.060116208
35	6243573	251168.5	5742374	6744772	0.919725612	1.080274388
36	6260131	290224.9	5680996	6839265	0.907488358	1.092511483
37	7849958	328896.6	7193655	8506261	0.916394075	1.083605925
38	5736501	173237.7	5390811	6082192	0.939738527	1.060261647
39	5915380	205874.8	5504564	6326197	0.930551207	1.069448962
40	5717169	204168.4	5309757	6124580	0.928738857	1.071260968
41	6080536	175565.9	5730200	6430873	0.942384027	1.057616138
42	6324489	235510.5	5854535	6794443	0.925692969	1.074307031
43	4277387	361103.4	3556816	4997958	0.831539442	1.168460558
44	5471038	205605.1	5060759	5881316	0.925008929	1.074990888
45	5471561	204443.1	5063601	5879521	0.925439925	1.074560075

Table 17: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	%OF LOWEST PRED.COST	%OF HIGHEST PRED.COST
46	5658019	205089.6	5248769	6067269	0.92766903	1.07233097
47	5796725	213509.2	5370674	6222776	0.92650143	1.07349857
48	6339934	176714.6	5987305	6692562	0.944379705	1.055620137
49	6435741	162348.3	6111780	6759701	0.949662207	1.050337638
50	5541060	171043.3	5199749	5882372	0.938403302	1.061596879
51	6386980	271904.6	5844403	6929557	0.915049523	1.084950477
52	9118063	449490.1	8221119	10015006	0.901629984	1.098369906
53	6918490	282557.1	6354656	7482324	0.918503315	1.081496685
54	6249920	340081.2	5571299	6928542	0.89141925	1.10858091
55	5664994	389091.7	4888574	6441414	0.86294425	1.13705575
56	4820565	197840.5	4425780	5215349	0.918103998	1.081895794
57	5683536	198050.3	5288333	6078739	0.930465295	1.069534705
58	6108130	180739.0	5747471	6468789	0.940954269	1.059045731
59	7198954	375795.7	6449066	7948843	0.895833756	1.104166383
60	5347346	225859.0	4896652	5798041	0.91571632	1.084283867
61	7480650	352658.5	6776931	8184369	0.905928094	1.094071906
62	5536700	252811.5	5032223	6041178	0.908884895	1.091115285
63	5514130	183877.1	5147208	5881051	0.933457862	1.066541957
64	6216423	228030.6	5761395	6671451	0.926802278	1.073197722
65	5930842	267760.4	5396535	6465150	0.909910431	1.090089738
66	6034389	214128.3	5607103	6461675	0.929191506	1.070808494
67	6125461	169081.2	5788065	6462858	0.944919084	1.055081079
68	5890264	267279.5	5356916	6423612	0.909452615	1.090547385

Table 17: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	%OF LOWEST PRED.COST	%OF HIGHEST PRED.COST
69	6284425	242598.0	5800328	6768521	0.922968768	1.077031073
70	5961545	213454.0	5535604	6387486	0.928551911	1.071448089
71	7441045	371767.9	6699194	8182896	0.900302847	1.099697153
72	4802506	243788.2	4316034	5288978	0.898704551	1.101295449
73	6142855	255961.6	5632092	6653619	0.916852506	1.083147657
74	6255833	163520.8	5929533	6582134	0.947840679	1.052159481
75	6051994	172645.7	5707485	6396503	0.943075125	1.056924875
76	5992580	154999.7	5683283	6301877	0.948386672	1.051613328
77	6612903	346607.0	5921260	7304547	0.895410079	1.104590072

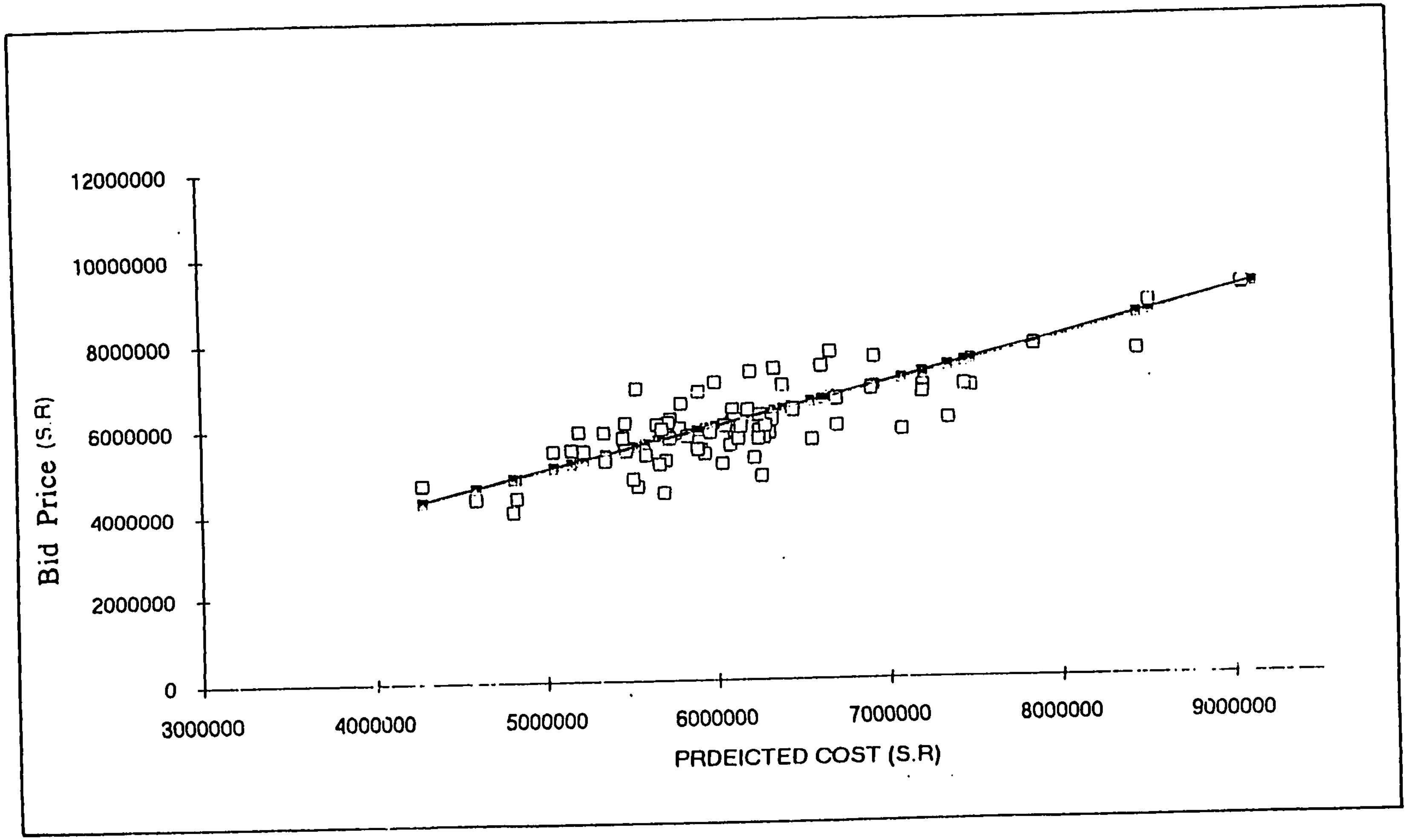

Figure 13 : Regression Line for Design Model No.2

4.3.3 School Design Model # 3

4.3.3.1. Determining the probability distribution parameters

In order to develop a statistical cost estimate model for school design model # 3, a sample of bid prices for 30 schools submitted to the Ministry of Education in 1994 was used. The general form of the potential regression equation is as follows:

$$\textbf{Predicted Cost} = \beta_0 + \beta_1 (X_1) + \beta_2 (X_2) + \beta_3 (X_3) + \ldots \beta_n = (X_n)$$

where X_1, X_2, X_3 and X_n are the independent variables used as concrete area, number of class rooms, bearing capacity of soil ...etc. The following assumptions are made:

1. The construction cost is a linear function of the independent variables.
2. No relation exists between the independent variables.

The validation of the first assumption was checked by analyzing and studying the residual plot, while the validation of the second assumption was performed by studying the correlation matrix of all the dependent variables shown in Appendix B. The linear least square regression was performed using the step wise procedure available in the **SAS** package to calculate the regression coefficients, β_k, and the variance of estimates, σE^2, of the construction cost regression equation.

After examining the correlation matrix, the variables found with high correlation factors (IRI > 0.5) are as follows:

Number of classrooms	-	Execution time
Execution time	-	Level of school
Concrete area	-	Execution time
Concrete area	-	Level of school
Concrete area	-	Number of technical laboratories
Length of retaining wall	-	Quantity of backfilling material
Level of school	-	Number of technical laboratories.
Concrete area	-	Number of class rooms

*Execution time, Length of retaining walls, Number of technical laboratories and Concrete area were dropped.

The different regression models were developed to investigate which set of variables produces the needed model with higher R^2. This analysis indicated that the following produce the highest R^2.

1. Length of fence
2. Number of classrooms
3. Availability of soil improvement
4. Bearing capacity of the soil
5. Level of school
6. Quantity of backfilling material
7. Quantity of excavated material

Therefore, the potential form for the general potential regression equation for the construction cost of school projects is as follows:

CC = β_0 + β_1 (Length of fence) + β_2 (Availability of soil improvement) + β_3 (Bearing capacity of soil)+β_4 (Level of school)+ β_5 (Quantity of backfilling materials)+ β_6 (Number of classrooms) + β_7 (quantity of excavation materials)

Where:

CC = School Construction Cost

β_2 = Regression constant

β_k = Factor weights, $K = 1,2,3,...7$

The linear least squares regression was performed using the step wise procedure available in the **SAS** package to calculate the regression coefficients, $\boldsymbol{\beta k}$, and the variance of estimates, σE^2, of the construction cost regression equation (Table 18). The regression analysis correlates the 30 schools' construction cost to the six potential factors. The R^2 value for the regression is 89.03 percent. The six factors explain 89.03 percent of the variation in the schools' construction cost.

Since it is risky to consider some of the variables which have a low significance level such as the bearing capacity, soil improvement and the number of class rooms, it was decided to go back among the different statistical models developed using the step wise procedure until the significance level of the selected parameters became higher than 95%. The selected statistical model is presented in Table 19.

TABLE 18 : REGRESSION COEFFICIENTS AND THEIR SIGNIFICANCE LEVEL FOR SCHOOL DESIGN MODEL # 3

Regression Equation :-

$CC = \beta_0 + \beta_1$ (# of class rooms) $+ \beta_2$ (excavation) $+ \beta_3$ (bearing capacity) $+ \beta_4$ (soil improvement) $+ \beta_5$ (level of scholl) $+ \beta_6$ (fill)

Coefficient of Multiple Determination $R^2 = 0.8903$

Multiple Correlation Coefficient R = 0.9435

Source Of Variation	Degree of Freedom	Sum Of Squares	Mean Squares	F Ratio	Significant At
Regression	6	38703164513706	6450527418951	31.14	99.99%
Residual	23	4764350517010	207145674652	-	-

Variable	Regression Coefficient	Standard Error of Estimate σ_E	Variance of Estimate σ_E^2	T-Ratio	Significant At
Constant	3579709.1	806699.7	6.50764406E11	4.43	99.98%
# of Class Rooms	24350.53	21700.3	470903020.1	1.12	72.66%
Excavation	102.1	47.95	2299.2	2.18	95.58%
Bearing Capacity	-270840.7	319719.3	1.022204308E11	0.848	59.43%
Soil Improvement	245017.3	318260.3	1.012896186E11	0.768	55.08%
Level of School	1025926.3	98554.8	9713048603	10.41	99.99%
Fill	133.41	30.15	909.022	4.45	99.98%

TABLE 19 : REGRESSION COEFFICIENTS AND THEIR SIGNIFICANCE LEVEL FOR SCHOOL DESIGN MODEL # 3

Regression Equation :-
$CC = \beta_0 + \beta_1$ (excavation) + β_2 (level of school) + β_3 (fill)

Coefficient of Multiple Determination $R^2 = 0.8713$

Multiple Correlation Coefficient R = 0.9334

Source Of Variation	Degree of Freedom	Sum Of Squares	Mean Squares	F Ratio	Significant At
Regression	3	37875762495863	12625254165288	58.7	99.99%
Residual	26	5591752534854.1	215067405186	-	-

Variable	Regression Coefficient	Standard Error of Estimate σ_E	Variance of Estimate σ_E^2	T-Ratio	Significant At
Constant	3437070	239199	5.72161616E10	14.36	99.99%
Excavtion	142.45	42.31	1790.13	3.36	99.76%
Level of school	1012044.08	91059.3	8291796116	11.11	99.99%
Fill	136.18	30.21	912.6441	4.5	99.99%

Since the confidence level for the selected model represented by its significance level is considered very high (99.99%), it can be concluded that the model is statistically significant. Regarding the parameters themselves, it can be seen that their confidence level is also very high (99.76% - 99.99%) and we can conclude that all the above parameters will be considered in the selected linear statistical model.

Thus, the linear statistical model for school design model number 3 built by the Ministry of Education in Saudi Arabia is as follows:

$$\textbf{Predicted Cost} = \mathbf{3437070.63 + 142.4577\ (excav) + 1012044.8\ (level) + 136.186\ (fill)}$$

where:

Excav	=	Quantity of excavated material
Level	=	Level of school
Fill	=	Quantity of backfilling material

As can be seen from the above statistical model, three different independent variables significantly contribute to the cost of boys schools built as per school design model number 3. As in the case of the general model and model number 2, these independent variables are classified into superstructure and substructures categories. The superstructure category includes the level of school, while the substructures category includes the quantity of excavated and backfilling materials.

Among all contributing factors, the greatest impact on the cost of school projects built as per this design model was found to be caused by the level of school factor. The level factor is really very important since it determines the availability and number of technical laboratories that are available in the school. As seen before, the presence of technical laboratories will increase the cost of projects due to their special characteristics and building requirements. The effect of this factor was found to be very high so that it increases the total cost of the project by 17.6% for every one more level. In other words, the construction cost of an intermediate school will be 17.6% more if compared to the construction cost of an elementary level school. The same also applies to the secondary level school if compared to the intermediate level school.

It was found from the interviews conducted with the engineers in the project studying department that the foundation for school design model number 3 is designed to carry 3 floors (i.e. 12 or 20 or 22 classrooms) regardless of the actual number of floors in the building. In other words the school with 12 classrooms is constructed using the same footing dimensions as a school with 20 or 22 classrooms and this is the reason for finding the number of classrooms, bearing capacity and availability of soil improvement not affecting the cost of the project as expressed in the developed model.

It was noticed from the collected data that 50% of the schools built as per design model number 3 had 12 classrooms only (i.e. 2 floors). Since this high percentage was built using foundation dimensions of 3 floors (20-22 classrooms) rather than 2 floors, it is thought that the extra cost paid for this over-design of footing dimensions and the high impact of the level of school on this specific design model made up the majority of the cost devoted to them, which of course reduced the importance of other factors such as length of fence and number of classrooms in expressing the total cost of the project.

Table 20: The Average Effect of Increase in one of the Parameters on the Overall Construction Cost of School Project for school design model number 3.

Parameter	Amount of Increase	Percent of Increase
Quantity of excavated material	$100m^3$	0.25%
Level of school	one more	17.6%
Quantity of backfilling	$100m^3$	0.24%

4.3.3.2 Testing of Assumptions

When a regression model is applied in practice, one cannot usually be sure in advance that the model implied is appropriate for the situation in hand. Consequently, the model considered needs to be checked for suitability and its agreement with the earlier assumptions made to develop it. As stated earlier, the following assumption were made:

i The regression function is linear

ii The distribution is normal

iii There is a constant variance over all schools construction cost.

In the following paragraphs, the above mentioned assumptions will be verified and checked.

4.3.3.2.1 Linearity Assumption

In order to test the developed model for its aptness and linearity, the residual plot was analyzed and studied. This graph was obtained by plotting the residuals against the fitted values (Figure 14). If the specified regression model is linear, the residuals will tend to scatter at random around the zero line when plotted against the fitted values. As can be seen in Figure 14 there is no pattern of systematic departure and all points are scattered normally around the zero line. This shows that the selected statistical model for all design models built by the Ministry of Education is linear and apt.

4.3.3.2.2 Normality and Constant Variance Assumptions

In order to check the normality assumption, and to be sure that a constant variance exists over all schools construction cost, the lack of normality test was performed. In this test, the standardized residuals were calculated. If the distribution was normal with the constant variance and the sample size is small (less than or equal 30 observations), the standardized residuals will tend to follow the standard normal distribution (figure 15). Hence, about half of the standardized residuals should be positive and half negative, About 50% of the standardized residual distribution should fall between t (0.25: n -2) and t (0.75: n -2) (i.e. $\mu \pm 1\sigma$) and at least 95% of the area under the normal distribution should be contained within t (0.025: n-2) and t (0.975: n-2) (i.e. $\mu \pm 2\sigma$) (Netir, Warerman and Whitemore, 1978).

In order to check if the developed model complies with the above mentioned criteria of normal distribution, the standardize residuals were calculated and listed in table 21.

Based on the analysis that was conducted on the results listed in table 21, the following was found:

t	(0.25 : 28)	=	- 0.683
t	(0.75 : 28)	=	+ 0.683
t	(0.025 : 28)	=	- 2.048
t	(0.975 : 28)	=	+2.048

$-0.683 < 96.4\ \%$ of the distribution < 0.683

96.6% of the distribution are negative

$-2.048 < 100\%$ of the distribution < 2.048

Since the above analysis complies with the different test criteria, then it can be concluded that this analysis provides no evidence of any departure form normality and the variance is constant over all school construction cost.

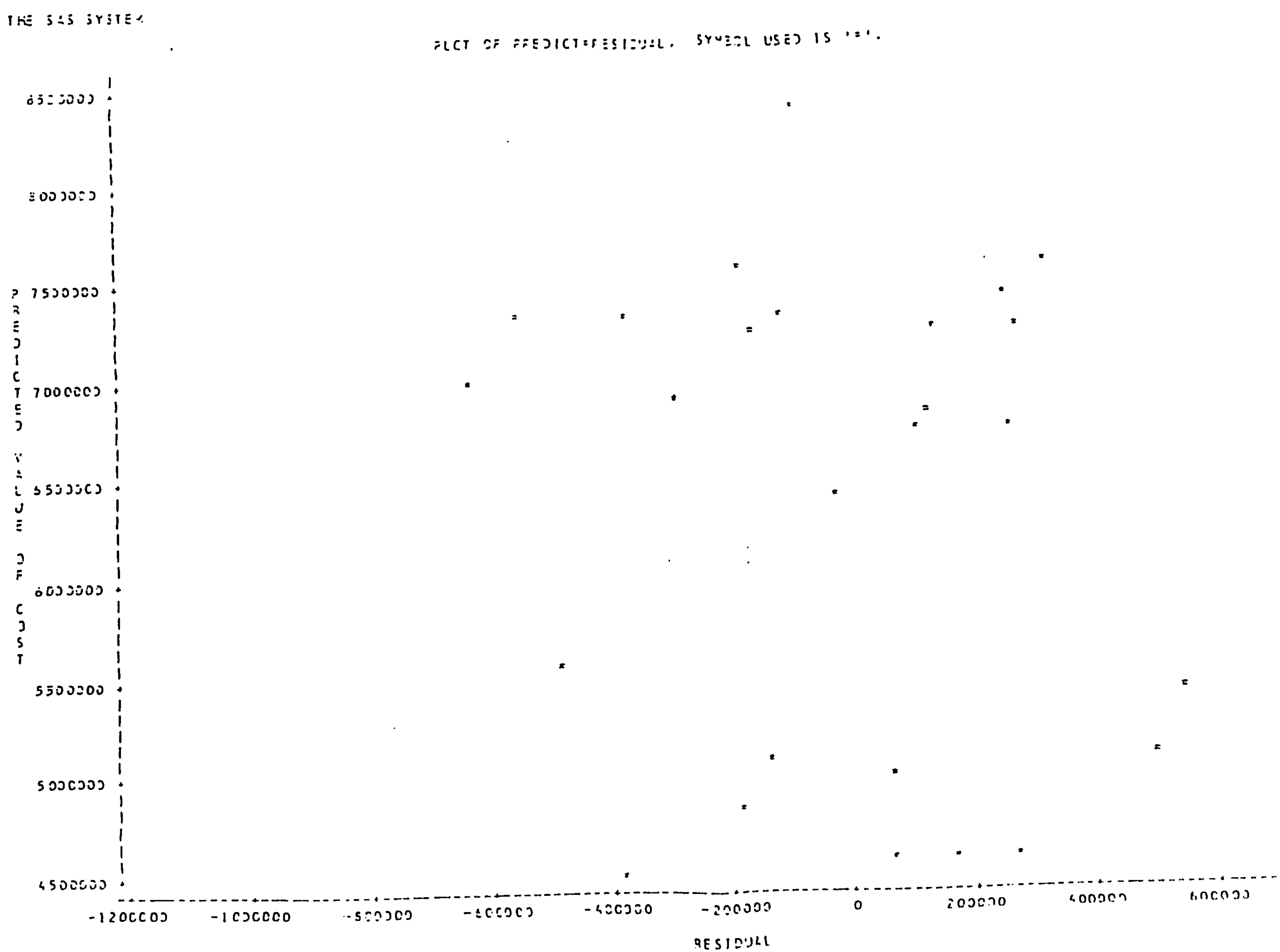

Figure 14: Residual Plot for School Design Model No.3

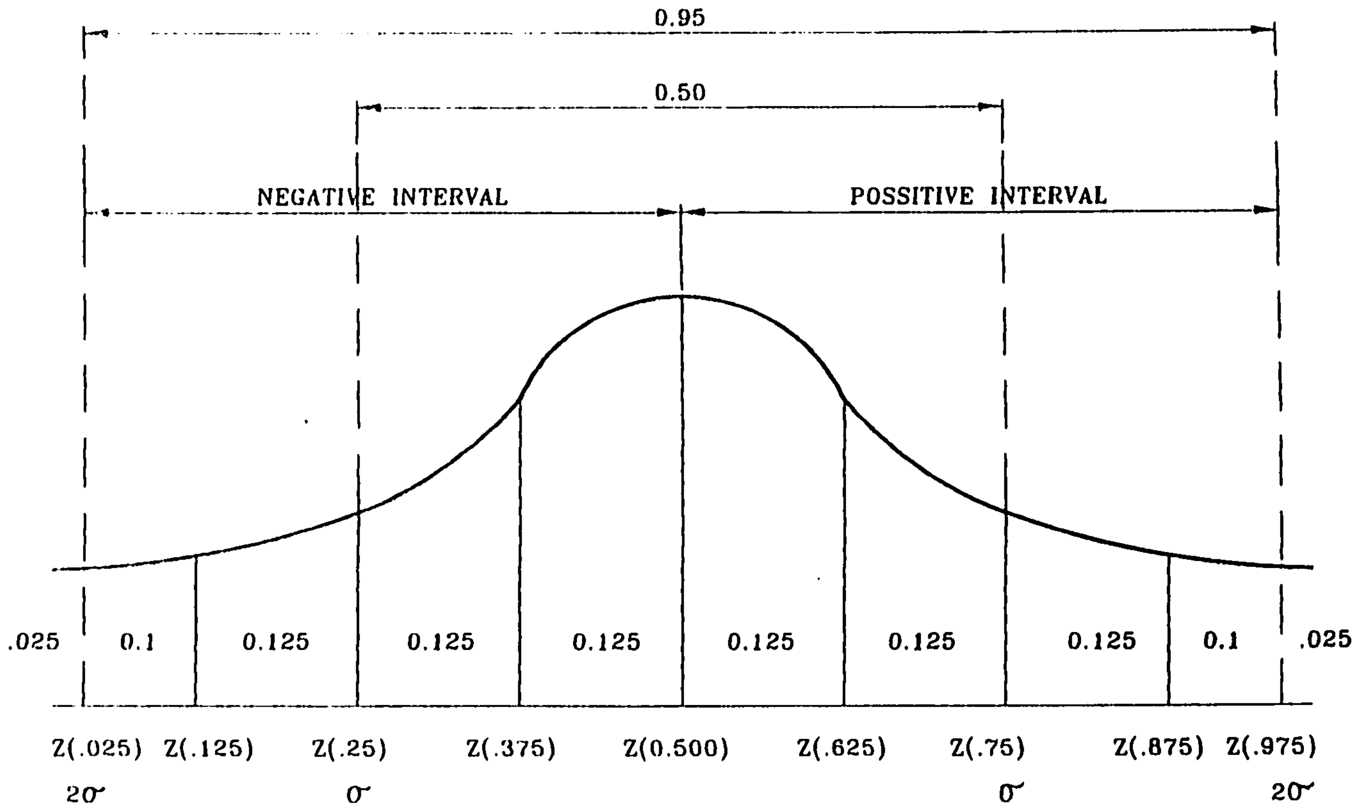

FIG:15. PARTITIONING OF A NORMAL DISTRIBUTION CURVE INTO 8 EQUAL INTERVALS FOR SMALL SAMPLE SIZE.

TABLE 21 : RESIDUAL TABLE FOR STATISTICAL MODEL # 3

NO	MODEL	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
1	3	5553587	5054514	499073	0.340311773
2	3	4642305	4817680	-175375	-0.119586067
3	3	6374790	6387146	-12356	-0.008425405
4	3	6901734	6781588	120146	0.081926087
5	3	4961254	6134809	-1173555	-0.800232796
6	3	6791200	7349499	-558299	-0.380697258
7	3	7883848	7558766	325082	0.221669439
8	3	7229710	7334998	-105288	-0.071794599
9	3	4863248	4594467	268781	0.183278475
10	3	7007962	6738155	269807	0.183978092
11	3	6941803	6806831	134972	0.092035755
12	3	6617453	6900412	-282959	-0.19294628
13	3	7538941	7249304	289637	0.197499927
14	3	7698578	7436691	261887	0.178577541
15	3	7566002	7286984	279018	0.190258961
16	3	5076755	5561825	-485070	-0.330763298
17	3	8307439	8378742	-71303	-0.048620643
18	3	7440296	7613551	-173255	-0.118140465
19	3	7935874	7282693	653181	0.445396132
20	3	4093342	4477052	-383710	-0.261647154
21	3	4942961	5081637	-138676	-0.094561468
22	3	6960810	7334281	-373471	-0.254665306

Table 21: Continued

NO	MODEL	COST	PREDITED	RESIDUL	STANDARDISED RESIDUL
23	3	6390231	7031796	-641565	-0.437475324
24	3	7408094	6788791	619303	0.422295139
25	3	7078125	7231517	-153392	-0.104596128
26	3	5059463	4985204	74259	0.050636304
27	3	6008365	5451992	556373	0.379383942
28	3	4645076	4585526	59550	0.040606416
29	3	4724756	4563368	161388	0.110048503
30	3	7389708	7233891	155817	0.106249706

4.3.3.3 Validity of the Model:

In order to test or to verify the validity of the model, six schools were selected randomly, excluded from the data base and withheld for checking the predictability of the developed model. The reliability of the model's predictability was examined by comparing the bid price with the model estimates (Table 22). As can be seen from table 22 that the develop model tends to over estimate the project cost. On the other hand, this model also possesses an acceptable balanced behavior when it over or under estimates the project cost (i.e. the percent difference between the bid price and the predicted cost reached close values with positive and negative sign). This balanced behavior adds more credits to the reliability and predictability of the developed model.

Moreover, by considering the fact that the model is primarily designed for the purpose of preliminary estimating and the variation in predicted and bid price costs falls within the acceptable limits (-30%, + 30%), then it can be concluded that the developed model is the most representative realistic model for school design model # 3.

Table 22: A Comparison Between the Bid Price and the Predicted Cost.

Project Name	Predicted Cost	Bid Price	Percent Difference
Talam*	5408054*	6337045	-15%
Jaber Bin Hayan	4975711	4784654	+ 4%
Khamis Motir	7641664	7501456	+ 2%
Tahfiz Al-Quran	6901643	6234213	+ 10%
Al-Bada	5918947	4954036	+ 19.5%
Safwa Bin Qadama	6013389	5472344	+ 10%

* Quantity of excavated materials = 900
Level of school = 1
Quantity of backfilling materials = 6100
Predicted Cost = 3437070.63 + 142.45 (900) + 1012044 (1)
+ 136.186 (6100)
= 5408054 (SR)

4.3.3.4 Predictability of the Model

In order to compare the predictability of the developed model against the estimated cost obtained by using the current estimating system of the Ministry of Education, a sample of schools was randomly selected. As can be seen in Table 23, the estimated cost using the developed model was greatly improved. This is thought to be due to the involvement of several influencing variables in the statistical model rather than only one variable (concrete area) as in the current estimating system used by the Ministry of Education. Another reason that might be behind the improvement obtained in the estimated cost is the great importance given to the potential cost factors in any construction project, such as the site conditions, or to the special factors which distinguish school projects from any other projects, such as its level (i.e. elementary, intermediate or secondary) and the available number of technical laboratories.

Table 23: Comparison Between the Predicted Cost Using the Developed Regression Model # 3 and the Cost Estimate Prepared by the Ministry of Education

Project Name	Design Model	Bid Price (1)	Predicted Cost Using the Regression Model (2)	Estimated Cost Using the Ministry's System(3)	Percentage Difference Between 1 & 2	Percentage Difference Between 1 & 3
JABER BIN HAYAN	3	4784654	4975711	5392400	+4%	+12%
KHAMIS NOTIN	3	7501456	7641664	5796000	+4%	-22%
TAHFIZ AL-QURAN	3	6234213	6901643	5796000	+10%	-8%
AL-BADA	3	4954036	5918947	3956750	+19.%	-20%
SAFWA BIN QUADAMA	3	5472344	6013389	3956750	+10%	-27%

4.3.3.5 Confidence Level of the Model

Confidence level can be interpreted as the probability of the estimate being correct. As being known, a good estimate implies a small length of confidence interval at high level of confidence. Thus, to measure the interval at which our estimate is correct, the upper and lower confidence limit were calculated at 95% confidence level. Moreover, the percentage of the upper and lower limits of the predicted cost with respect to their mean were calculated to have an estimate of how the predicted cost may vary within the above specified level of confidence. In table 24, the predicted cost and their statistical analysis as discussed above are presented.

By referring to table 24, it can be noticed that at 95% of confidence level the predicted costs may vary with ±11% only from their mean value. This low confidence interval is considered as an indication for the goodness and reliability of the estimated cost obtained from the developed statistically model.

TABLE 24 : CONFIDENCE LEVEL FOR REGRESSION MODEL # 3

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	%OF LOWEST PRED.COST	%OF HIGHEST PRED.COST
1	5054514	204490.7	4631493	5477535	0.916308274	1.083691726
2	4817680	165332.9	4475663	5159697	0.929007946	1.070992054
3	6387146	347949.8	5667356	7106935	0.887306475	1.112693369
4	6781588	170083.3	6429744	7133432	0.948117756	1.051882244
5	6134809	169512.2	5784147	6485472	0.9428406	1.057159563
6	7349499	175928.3	6985563	7713434	0.950481523	1.049518341
7	7558766	312533.4	6912241	8205291	0.914466859	1.085533141
8	7334998	170391.7	6982516	7687480	0.951945181	1.048054819
9	4594467	179194.3	4223776	4965159	0.919317954	1.080682264
10	6738155	160586.5	6405956	7070353	0.950698819	1.049301033
11	6806831	152971.1	6490386	7123276	0.953510672	1.046489328
12	6900412	181916.5	6524089	7276735	0.945463691	1.054536309
13	7249304	170482.6	6896633	7601974	0.951351054	1.048648808
14	7436691	221475.6	6978534	7894848	0.938392358	1.061607642
15	7286984	150142.3	6976391	7597577	0.957377016	1.042622984
16	5561825	288130.6	4965781	6157869	0.892833018	1.107166982
17	8378742	312943.8	7731368	9026115	0.922736134	1.077263747
18	7613551	226023.1	7145987	8081115	0.93858792	1.06141208
19	7282693	149499.5	6973430	7591956	0.957534527	1.042465473
20	4477052	175741.7	4113503	4840602	0.918797235	1.081202988
21	5081637	232437.7	4600803	5562470	0.905378129	1.094621674
22	7334281	235437.3	6847242	7821320	0.933594172	1.066405828

Table 24: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	%OF LOWEST PRED.COST	%OF HIGHEST PRED.COST
23	7031796	184555.7	6650013	7413578	0.945706189	1.054293668
24	6788791	155162.7	6467812	7109770	0.952719269	1.047280731
25	7231517	309895.1	6590450	7872584	0.911350966	1.088649034
26	4985204	198229.0	4575136	5395272	0.917742985	1.082257015
27	5451992	289438.7	4853243	6050742	0.890177939	1.109822245
28	4585526	169929.8	4234000	4937053	0.923340092	1.076660126
29	4563368	170205.5	4211271	4915465	0.922842734	1.077157266
30	7233891	315602.2	6581018	7886764	0.909748018	1.090251982

4.3.3.6 Fitting of Regression Line

As is known, the regression function represented by its plotted line provides a good fit to the sample data if it results in small residuals. So, to measure the degree of fitting of the developed model to the sample data, the residuals listed in Table 21 were carefully studied and analyzed. Moreover, the regression line was also plotted as seen in Figure 16. It was found that the bid price data represented by their plotted points were distributed on both sides of the fitted lines. Moreover, the dispersion was found to be small and in an average range of 4% of the actual cost. Hence, it can be concluded that the developed regression function provides a good fit to the sample data.

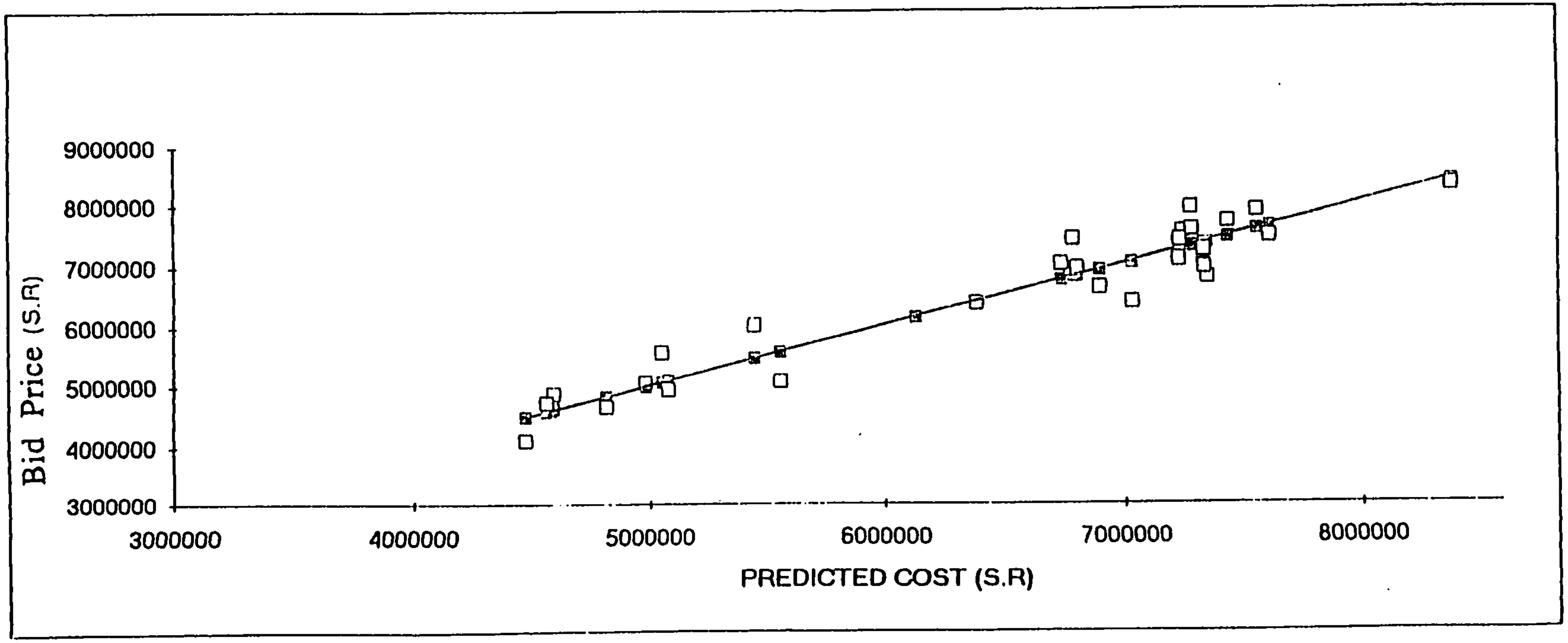

Figure 16: Regression Line for Design Model No. 3

4.3.4 School Design Model # 5

4.3.4.1 Determining the probability distribution parameters

In order to develop a statistical cost estimate model for school design model number 5, a sample of the bid prices for 20 schools submitted to the Ministry of Education in 1995 was used. The general form of the potential regression equation is as follows:

$$\textbf{Predicted Cost} = \beta_0 + \beta_1 (X_1) + \beta_2 (X_2) + b_3 (X_3) + \dots \beta n = (X_n)$$

where X_1, X_2, X_3 and X_n are the independent variables used as concrete area, number of classrooms, bearing capacity of soil, etc. The following assumptions are made:

1. The construction cost is a linear function of the independent variables.
2. No relation exists between the independent variables.

The validity of the first assumption was checked by analyzing and studying the residual plot, while the validation of the second assumption was performed by studying the correlation matrix of all the dependent variables shown in Appendix B. The linear least square regression was performed using the step wise procedure available in the **SAS** package to calculate the regression coefficients, β_k, and the variance of estimates σ_E^2, of the construction cost regression equation.

After examining the correlation matrix, the variables found with high correlation factors (IRI > 0.5) are as follows:

Length of retaining wall	-	Execution time
Execution time	-	Number of class rooms
Soil Improvement	-	Bearing Capacity
Level of school	-	Quantity of backfilling materials
Level of school	-	Number of technical laboratories
Number of technical laboratories	-	Quantity of backfilling material
Length of retaining wall	-	Quantity of excavated material
Execution time	-	Concrete area
Concrete area	-	Number of class rooms

*Fill, number of technical laboratories, Quantity of excavated material, Execution time, Concrete area were dropped.

The different regression models were developed to investigate which set of variables produces the needed model with higher R^2. This analysis indicated the following variables to produce the highest R^2.

1. Length of fence
2. Number of classrooms
3. Length of retaining walls
4. Availability of soil improvement
5. Level of school
6. Footing type
7. Bearing capacity of soil

Therefore, the potential form for the general potential regression equation for the. construction cost of school projects is as follows:

CC = β_0 + β_1 (Length of fence) + β_2 (number of classrooms) + β_3 (Length of retaining walls) + β_4 (Availability of soil improvement) + β_5 (Level of school) + β_6 (Foot-type) + β_7 (Bearing capacity of soil)

Where:

CC = School Construction Cost

β_0 = Regression constant

β_k = Factor weights, K = 1,2,3........7

The linear least squares regression was performed using the step wise procedure available in the **SAS** package to calculate the regression coefficients, βk, and the variance of estimates, σE^2, of the construction cost regression equation (Table 25). The regression analysis correlates the 20 schools' construction cost to the four potential factors. The R^2 value for the regression is 67.19 percent. The four factors explain 67.19 percent of the variation in the schools, construction cost.

Since it is risky to consider some of the variables which have a low significance level such as soil improvement, it was decided to go back among the different statistical models developed using the step wise procedure until the significance level of the selected parameters became higher than 95%. The selected statistical model is presented in Table 26.

TABLE 25 : REGRESSION COEFFICIENTS AND THEIR SIGNIFICANCE LEVEL FOR SCHOOL DESIGN MODEL # 5

Regression Equation :-

$CC = \beta_0 + \beta_1$ (# of class rooms) + β_2 (Length of retaining walls) + β_3 (Soil improvement) + β_4 (level of schools)

Coefficient of Multiple Determination $R^2 = 0.6719$

Multiple Correlation Coefficient R = 0.8196

Source Of Variation	Degree of Freedom	Sum Of Squares	Mean Squares	F Ratio	Significant At
Regression	4	15238067940316	3809516985078	7.68	99.86%
Residual	15	7439674623012	495978308200	-	-

Variable	Regression Coefficient	Standard Error of Estimate σ_E	Variance of Estimate σ_E2	T-Ratio	Significant At
Constant	2208307.4	1233917.7	1.52255289E12	1.78	99.63%
# of Class Rooms	77325.917	40280.65	1622530764	1.92	92.59%
Length of retaining wall	8858.43	2176.79	4738414.7	4.06	99.90%
Soil Improvement	-407212.4	564674.35	3.18857121E11	0.721	51.81%
Level of School	988483.6	542648.6	2.944675031E11	1.82	91.15%

TABLE 26 : REGRESSION COEFFICIENTS AND THEIR SIGNIFICANCE LEVEL FOR SCHOOL DESIGN MODEL # 5

Regression Equation :-
$CC = \beta_0 + \beta_1$ (# of class rooms) + β_2 (Length of retaining walls) + β_3 (Level of schools)

Coefficient of Multiple Determination $R^2 = 0.6605$

Multiple Correlation Coefficient $R = 0.8127$

Source Of Variation	Degree of Freedom	Sum Of Squares	Mean Squares	F Ratio	Significant At
Regression	3	14980133869786	4993377956595	10.38	99.95%
Residual	16	7697608693542	481100543346	-	-

Variable	Regression Coefficient	Standard Error of Estimate σ_E	Variance of Estimate σ_E^2	T-Ratio	Significant At
Constant	1856791.5	1116427.3	1.246409916E12	1.664	88.43%
# of Class Rooms	87340.7	37239.46	1386777381	2.34	96.78%
Length of retaining wall	8868.8	2143.8	4595878.44	4.13	99.92%
Level of School	1058720.4	525768.9	2.764329362E11	2.01	93.88%

Since the confidence level for the selected model represented by its significance level is considered very high (99.95%), it can be concluded that the model is statistically significant. Regarding the parameters themselves, it can be seen that their confidence level is also very high (88.43% - 99.92%) and we can conclude that all the above parameters will be considered in the selected linear statistical model.

The linear statistical model for school design model number 5 built by the Ministry of Education in Saudi Arabia is as follows:

Predicted Cost =1856791.5+87340.7 (rooms) +8868.8 (RW)
+ 1058720.4 (level)

where:

Room	=	number of classrooms
RW	=	length of retaining walls (m)
Level	=	level of school

As can be seen from the above statistical model, three different independent variables significantly contribute to the cost of boys schools built as per school design model number 5. Contrary to all other design models that are being used by the Ministry of Education, the construction cost for this model is dependent only on superstructure factors such as the number of classrooms, length of retaining walls and level of school.

It was found during the interviews conducted with the engineers in the project studying department that there was a need for a design of a school building to be constructed in the Makkah area which has the same number of classrooms but with less total area than any other available design models with the Ministry of Education. This need was due to the fact that land in Makkah is very precious and is very difficult to find large empty areas within the town. In order to fulfill these design requirements, school design model No. 5 was developed.

After examining the developed statistical cost estimate model, it can be seen that all factors related to the substructure such as soil improvement, and site preparation activities (i.e., excavation and backfilling) are not significant. This is due to the rocky soil in the Makkah area. It is known that this type of soil does not usually need improvement and it is very expensive to excavate. The bearing capacity in such areas is also high and this is the reason for using only one type of footing (isolated footing) and for finding this factor not contributing much to the statistical model.

Since large empty areas in the Makkah area not abundant, school design model number 5 expands vertically rather than horizontally. In other words, the minimum number of floors is 3 for 14 classrooms and it reaches up to 4 floors in case the of 28 classrooms. This vertical expansion in the design itself makes the cross sectional area of some structural elements such as columns and footings higher if compared to a school with a lower number of floors. The higher cross sectional area of these structural elements implies a higher quantity of reinforced concrete material used and accordingly higher cost. As the number of classrooms increases, the number of students also increases. The increase in the number of students will imply also an increase in the necessary supporting services such as toilets and their plumbing system, which of course will increase the cost of the school projects. All the above mentioned factors made the number of classrooms significantly affect the cost of the school project built as per design model number 5. This significant effect was found to add 1.58% to the total construction cost of the project if the number of classrooms is be increased by one .

Another factor found to contribute significantly to the cost of school projects built as per school design model number 5 is the length of retaining walls. The standard specifications of the Ministry of Education state that if the difference in level between outside and inside the school is one meter, then a retaining wall is needed. This close limit implies the need for retaining walls in some places especially when a great areas given to the project. As is known, the construction of a retaining wall is expensive compared to the construction of a fence made of hollow blocks and it will increase the cost of the project dramatically. The effect of this factor on the total construction cost of a school project was found to be 1.61% for every additional 10(m).

Table 27: The Average Effect of Increase in one of the Parameters on the Overall Construction Cost of School Project for school design model number 5.

Parameter	Amount of Increase	Percent of Increase
Number of classrooms	1	1.58%
Length of retaining walls	10(m)	1.61%
Level of school	one more	19 %

4.3.4.2 Testing of Assumption

When a regression model is applied in practice, one cannot usually be sure in advance that the model implied is appropriate for the situation in hand. Consequently, the model considered needs to be checked for suitability and its agreement with the earlier assumptions made to develop it. As stated earlier, the following assumption were made:

i The regression function is linear

ii The distribution is normal

iii There is a constant variance over all schools construction cost.

In the following paragraphs, the above mentioned assumptions will be verified and checked.

4.3.4.2.1 Linearity Assumption

In order to test the developed model for its aptness and linearity, the residual plot was analyzed and studied. This graph was obtained by plotting the residuals against the fitted values (Figure 17). If the specified regression model is linear, the residuals will tend to scatter at random around the zero line when plotted against the fitted values. As can be seen in Figure 17 there is no pattern of systematic departure and all points are scattered normally around the zero line. This shows that the selected statistical model for all design models built by the Ministry of Education is linear and apt.

4.3.4.2.2 Normality and Constant Variance Assumptions

In order to check the normality assumption, and to be sure that a constant variance exists over all schools construction cost, the lack of normality test was performed. In this test, the standardize residual was calculated. If the distribution was normal with the constant variance and the sample size is small (less than or equal 30 observations), the standardize residual will tend to follow the standard normal distribution figure 15. Hence, about half of the standardized residuals should be positive and half negative, About 50% of the standardized residual distribution should fall between t (0.25: n -2) and t (0.75: n -2) (i.e. $\mu \pm 1\sigma$) and at least 95% of the area under the normal distribution should be contained within t (0.025: n-2) and t (0.975: n-2) (i.e. $\mu \pm 2\sigma$) (Netir, Warerman and Whitemore, 1978).

In order to check if the developed model complies with the above mentioned criteria of normal distribution, the standardize residuals were calculated and listed in table 28.

Based on the analysis that was conducted on the results listed in table 28, the following is found:

t	(0.25 : 18)	=	- 0.688
t	(0.75 : 18)	=	+ 0.688
t	(0.025 : 18)	=	- 2.101
t	(0.975 : 18)	=	+2.101

- 0.683 < 50% of the distribution < 0.683

55% of the distribution are negative

-2.101 < 95% of the distribution < 2.101

Since the above analysis complies with the different test criteria, then it can be concluded that this analysis provides no evidence of any departure form normality and the variance is constant over all school construction cost.

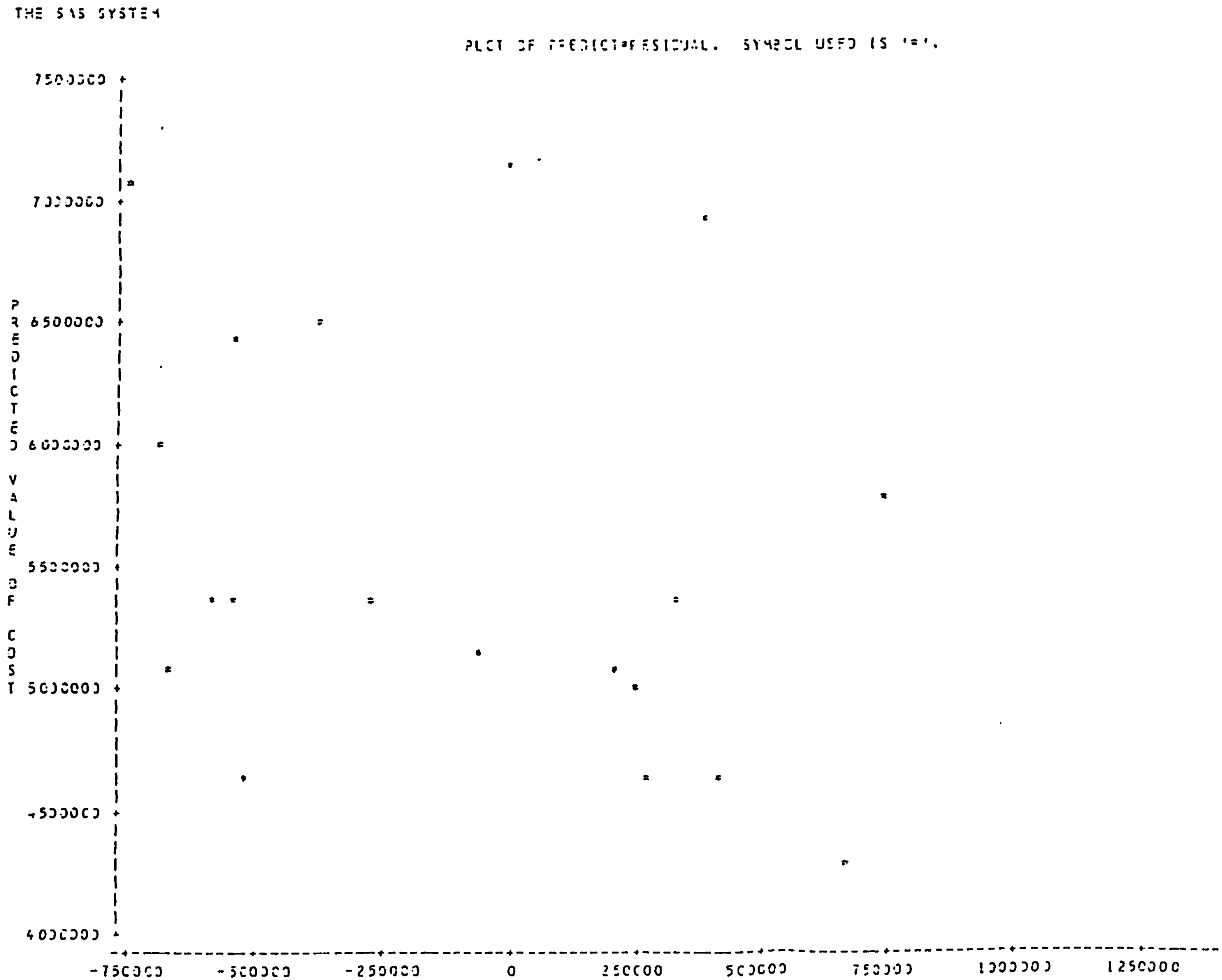

Figure 17: Residual Plot for Design Model No. 5

TABLE 28 : RESIDUL TABLE FOR STATISTICAL MODEL # 5

NO	MODEL	COST	PREDICTED	RESIDUL	STANDARDIZED RESIDUL
1	5	6559178	5809115	750063	1.081383882
2	5	5073788	4665980	407808	0.58794661
3	5	7176342	7177889	-1547	-0.002230347
4	5	5271682	5055751	215931	0.31131292
5	5	6100363	6469215	-368852	-0.531782807
6	5	5890731	6437666	-546935	-0.788529355
7	5	4134785	4665980	-531195	-0.765836618
8	5	4825280	5361911	-536631	-0.77367383
9	5	5093569	5361911	-268342	-0.386875121
10	5	4780251	5361911	-581660	-0.838593223
11	5	6299232	7049295	-750063	-1.081383882
12	5	4434368	5091460	-657092	-0.947345353
13	5	5688770	5361911	326859	0.471240488
14	5	5098395	5162052	-63657	-0.091775829
15	5	7344552	6947570	396982	0.572338505
16	5	4938182	4665980	272202	0.392440176
17	5	8541311	6823552	1717759	2.476534499
18	5	5275020	5029175	245845	0.354440654
19	5	5297435	5982001	-684566	-0.986955281
20	5	4915859	4258767	657092	0.947345353

4.3.4.3 Validity of the Model:

In order to test or to verify the validity of the model a number of projects were excluded from the data that were utilized in developing the model and withheld for checking the predictability of the developed model. A comparison between the predicted cost using the developed model and the bid price is presented in Table 29 as follows:

Table 29: A Comparison Between the Bid Price and Predicted Cost for school design model number 5

Project Name	Predicted Cost	Bid Price	Percent Difference
Ibn Maga*	6916988*	6659226	+ 3%
Ali Bin Aditalib	4574986	4939576	-8%
Tahfiz Al-Quran	6464041	6371614	+ 1%
Jafa	5631574	5702058	-2%

* Number of classrooms = 21
Length of retaining walls = 125 (m)
Level of school = 2

Predicted Cost = 1856791.54 + (87340.77 x 21) + (8868.8 x 125) + 1058720.47 (2)
= 6916988.65 (SR)

Considering the fact that the model is primarily designed for the purpose of preliminary estimating and, moreover, the variation in predicted and bid price falls within the acceptable limits (-30%, + 30%), then it can be concluded that the developed model is the most representative realistic model for school design model # 5.

4.3.4.4 Predictability of the Model

In order to compare the predictability of the developed model against the estimated cost obtained by using the current estimating system of the Ministry of Education, a sample of schools was randomly selected. As can be seen in Table 30, the estimated cost using the developed model was greatly improved. This is thought to be due to the involvement of several influencing variables in the statistical model rather than only one variable (concrete area) as in the current estimating system used by the Ministry of Education. Another reason that might be behind the improvement obtained in the estimated cost is the great importance given to the potential cost factors in any construction project, such as the site conditions, or to the special factors which distinguish school projects from any other projects, such as its level (i.e. elementary, intermediate or secondary) and the available number of technical laboratories.

Table 30: Comparison Between the Predicted Cost Using the Developed Regression Model # 5 and the Cost Estimate Prepared by the Ministry of Education

Project Name	Design Model	Bid Price (1)	Predicted Cost Using the Regression Model (2)	Estimated Cost Using the Ministry's System(3)	Percentage Difference Between 1 & 2	Percentage Difference Between 1 & 3
IBN- MAJA	5	6659226	6916988	6042750	+3%	-10%
ALI BIN ABITALIB	5	4939576	4574986	4311200	-8%	-12%
TAHFIZ AL-QURAN	5	6371614	6464041	4311200	+1%	-32%
JAFA	5	5702058	5631574	4438000	-2%	-22%

4.3.4.5 Confidence Level of the Model

Confidence level can be interpreted as the probability of the estimate being correct. As being known, a good estimate implies a small length of confidence interval at high level of confidence. Thus, to measure the interval at which our estimate is correct, the upper and lower confidence limit were calculated at 95% confidence level. Moreover, the percentage of the upper and lower limits of the predicted cost with respect to their mean were calculated to have an estimate of how the predicted cost may vary within the above specified level of confidence. In table 32, the predicted cost and their statistical analysis as discussed above are presented.

By referring back to table (32), it can be seen that the bid price may vary from the predicted value as listed in Table (31) below.

Table (31): Variation of Bid Price with Respect to the Predicted Cost for school design model number 5

Percent of the total observation points	Range of Variation at 95% confidence level
25%	±10%
60%	±15%
15%	±25%

From the above results, it can be concluded that at 95% confidence level, the difference between the predicted cost and their means for the majority of the observation points (85%) is below 15% while three observations only reached upto 25%. This low confidence interval for the majority of the data points is considered as a good indication for the goodness and reliability of estimated cost obtained from the developed statistical model.

4.3.4.6 Fitting of Regression Line

As is known, the regression function represented by its plotted line provides a good fit to the sample data if it results in small residuals. So, to measure the degree of fitting of the developed model to the sample data, the residuals listed in Table 28 were carefully studied and analyzed. Moreover, the regression function was also plotted as seen in Figure 18. It was found that the bid price data represented by their plotted points were distributed on both sides of the fitted lines. Moreover, the dispersion was found to be small and in an average range of 8 % of the actual cost. Hence, it can be concluded that the developed regression function provides a good fit to the sample data.

TABLE 32 : CONFIDENCE LEVEL FOR REGRESSION MODEL # 5

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
1	5809115	520775.8	4699107	6919122	0.808919603	1.191080225
2	4665980	286528.8	4055258	5276701	0.869111741	1.130888045
3	7177889	354995.0	6421236	7934543	0.894585581	1.105414559
4	5055751	266382.0	4487971	5623530	0.88769621	1.112303592
5	6469215	252191.4	5931682	7006748	0.916909084	1.083090916
6	6437666	423443.3	5535118	7340214	0.859801984	1.140198016
7	4665980	286528.8	4055258	5276701	0.869111741	1.130888045
8	5361911	295787.2	4731456	5992367	0.882419719	1.117580467
9	5361911	295787.2	4731456	5992367	0.882419719	1.117580467
10	5361911	295787.2	4731456	5992367	0.882419719	1.117580467
11	7049295	520775.8	5939288	8159303	0.842536452	1.15746369
12	5091460	508386.0	4007861	6175059	0.787173227	1.212826773
13	5361911	295787.2	4731456	5992367	0.882419719	1.117580467
14	5162052	266656.3	4593687	5730416	0.88989553	1.110104276
15	6947570	314772.7	6276648	7618492	0.903430696	1.096569304
16	4665980	286528.8	4055258	5276701	0.869111741	1.130888045
17	6823552	295346.4	6194036	7453068	0.90774365	1.09225635
18	5029175	266713.4	4460689	5597662	0.886962375	1.113037824
19	5982001	236366.1	5478199	6485804	0.915780355	1.084219812
20	4258767	508386.0	3175168	5342366	0.745560393	1.254439607

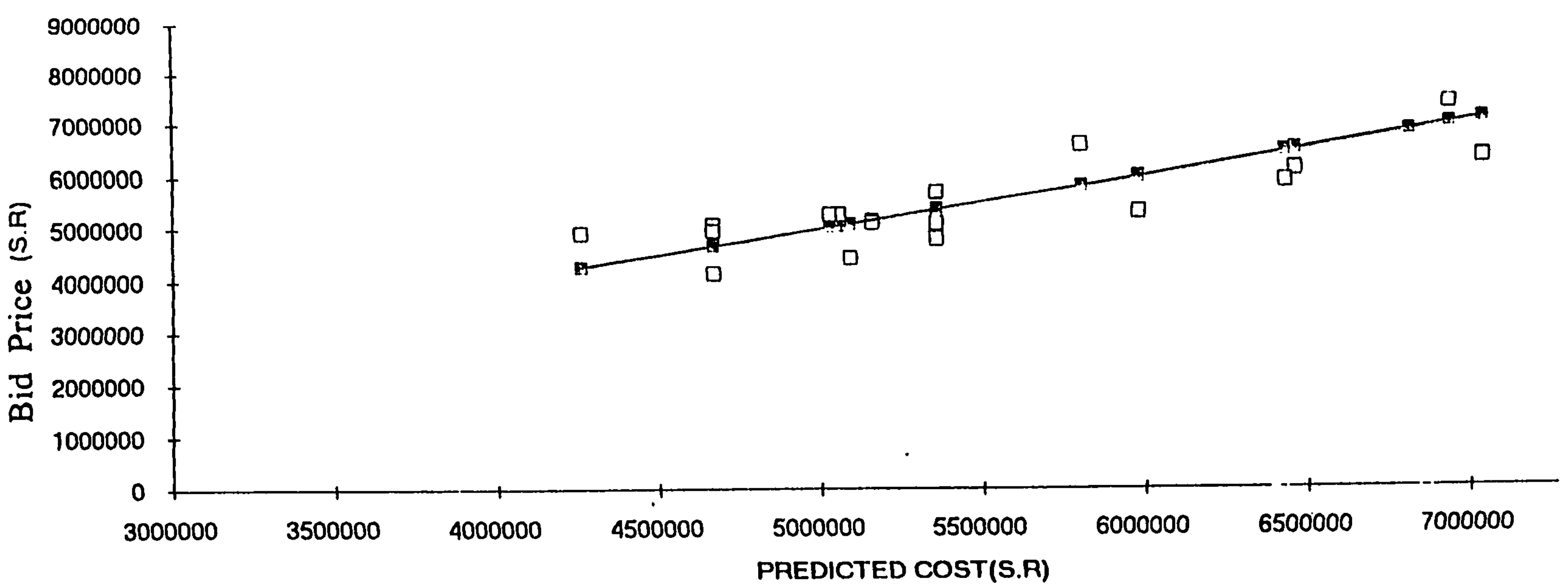

Figure 18: Regression Line for Design Model No. 5

4.3.5 *School Design Model # 11*

4.3.5.1. *Determining the probability distribution parameter*

In order to develop a statistical cost estimate model for school design model # 11, a sample of bid prices for 39 schools submitted to the Ministry of Education in 1994 was used. The general form of the potential regression equation is as follows:

$$\textbf{Predicted Cost} = \beta_0 + \beta_1 (X_1) + \beta_2 (X_2) + \beta_3 (X_3) + \dots \beta_n = (X_n)$$

where X_1, X_2, X_3 and X_n are the independent variables used as concrete area, number of classrooms, bearing capacity of soil, etc. The following assumptions are made:

1. The construction cost is a linear function of the independent variables.
2. No relation exists between the independent variables.

The validity of the first assumption was checked by analyzing and studying the residual plot, while the validation of the second assumption was performed by studying the correlation matrix of all the dependent variables shown in Appendix B. The linear least square regression was performed using the step wise procedure available in the **SAS** package to calculate the regression coefficients, β_k, and the variance of estimate, σ_E^2, of the construction cost regression equation.

After examining the correlation matrix, the variables found with high correlation factors (IRI > 0.5) are as follows:

Level of school	-	Number of class rooms
Soil improvement	-	Bearing capacity
Concrete area	-	Number of class rooms
Number of technical laboratories	-	Number of class rooms
Number of technical laboratories	-	Level of school
Length of retaining walls	-	Quantity of backfilling materials

* Level of school, soil improvement, Number of technical laboratories, concrete area and length of retaining walls were dropped.

The different regression models were developed to investigate which set of variables produces the needed model with higher R^2. This analysis indicated the following variables to produce the highest R^2.

1. Number of classrooms

2. Length of fence

3 . Execution time

4. Foot type

5. Bearing Capacity

6. Quantity of excavated materials

7. Quantity of backfilling materials

Therefore, the potential form for the general potential regression equation for the construction cost of school projects is as follows:

CC = β_0 + β_1 (Number of class rooms) + β_2 (length of fence) + β_3 (execution time) + β_4 (Foot type) + β_5 (Bearing capacity of soil) + β_6 (Quantity of excavated material) + β_7 (Quantity of backfilling).

Where:

CC	=	School Construction Cost
β_0	=	Regression constant
β_k	=	Factor weights, K = 1,2,3,...7

The linear least regression was performed using the step wise procedure available in the **SAS** package to calculate the regression coefficient, β_k, and variance of estimates, $\sigma_E{}^2$ of the construction cost regression equation (Table 33). The regression analysis correlates the 39 schools' construction cost to the five potential factors. The R^2 value for the regression is 56.57 percent. The five factors explain 56.57 percent of the variation in the schools' construction cost.

Since it is risky to consider some of the variables which have a low significance level such as excavation capacity, it was decided to go back among the different statistical models developed using the step wise procedure until the significance level of the selected parameters became higher than 95%. The selected statistical model is presented in Table 34.

Since the confidence level for the selected model represented by its significance level is considered very high (99.99%), it can be concluded that the model is statistically significant. Regarding the parameters themselves, it can be seen that their confidence level is also very high (96.76% - 99.75%) and we can conclude that all the above parameters will be considered in the selected linear statistical model.

TABLE 33 : REGRESSION COEFFICIENTS AND THEIR SIGNIFICANCE LEVEL FOR SCHOOL DESIGN MODEL # 11

Regression Equation :-

$CC = \beta_0 + \beta_1$ (length of fence) + β_2 (footing type) + β_3 (# of class rooms) + β_4 (excavation) + β_5 (fill)

Coefficient of Multiple Determination $R^2 = 0.5657$

Multiple Correlation Coefficient $R = 0.7521$

Source Of Variation	Degree of Freedom	Sum Of Squares	Mean Squares	F Ratio	Significant At
Regression	5	31181924887876	6236384977575	8.6	99.99%
Residual	33	23934473602542	725287078864	-	-

Variable	Regression Coefficient	Standard Error of Estimate σ_E	Variance of Estimate σ_E^2	T-Ratio	Significant At
Constant	4447235.12	1360730.8	1.85158831E12	3.26	99.75%
Length of fence	5644.65	3162.45	10001090	1.78	91.65%
Footing Type	1225276.8	489407.97	2.395201611E11	2.50	98.26%
# of class rooms	147081.05	46644.8	2175737367	3.15	99.66%
Excavation	35.02	32.94	1085.04	1.06	70.75%
Fill	20.58	15.96	254.72	1.28	79.62%

TABLE 34 : REGRESSION COEFFICIENTS AND THEIR SIGNIFICANCE LEVEL FOR SCHOOL DESIGN MODEL # 11

Regression Equation :-
$CC = \beta_0 + \beta_1$ (length of fence) + β_2 (footing type) + β_3 (# of class rooms)

Coefficient of Multiple Determination $R^2 = 0.5267$

Multiple Correlation Coefficient R = 0.7257

Source Of Variation	Degree of Freedom	Sum Of Squares	Mean Squares	F Ratio	Significant At
Regression	3	29033013260368	9677671086789	12.99	99.99%
Residual	35	26083385230050	745239578001	-	-

Variable	Regression Coefficient	Standard Error of Estimate σ_E	Variance of Estimate σ_E^2	T-Ratio	Significant At
Constant	4194245.8	1366541.1	1.867434578E12	3.06	99.59%
Length of fence	6779.4	3042.9	9259249.41	2.22	96.76%
Footing Type	1229399.4	494471.6	2.445021632E11	2.48	98.22%
# of class rooms	151488.1	46420.7	2154881388	3.26	99.75%

The linear statistical model for school design model # 11 built by the Ministry of Education in Saudi Arabia is as follows:

Predicted Cost = 4194245.8 + 6779.4 (fence) + 1229399.4 (foot type) + 151488.1 (rooms)

where:

fence	=	length of fence (m)
foot type	=	type of footing
rooms	=	Number of class rooms

It can be seen from the above statistical model that three different independent variables significantly contribute to the cost of boys schools built as per school design model number 11. As per the majority of the different statistical models developed earlier, these independent variables are classified in to superstructure and substructure factors. The supper structure factors are those including length of fence and level of school, while the substructure variable is the type of footing. As can be seen the length of fence was found to contribute to the construction cost of schools built as per model number 11. It was noticed from the data collected from the archive of the Ministry of Education that this design model usually used for larger sites than all other design models due to its large concrete area. These large sites usually have a longer fence and accordingly higher cost. The longer fence might also entail the use of longer strip footing if needed, which of course would add dramatically to the total cost of the project. The effect of this factor on the total cost of the project was found to be 1.1% for every 10 additional meters.

It was found during the interviews conducted with the engineers in the project studying department that this model is distinguished because of its large concrete area. This large concrete area means that the majority of cost consumption is devoted for building the class rooms and all concrete shaded supporting facilities such as the closed gymnasium and teachers' conference rooms, etc. This combination of large concrete area and large fence means the majority of the cost is spent on the superstructure facilities rather than the substructure facilities such as soil improvement and site preparation activities (i.e. excavation and backfilling). The effect of the number of classrooms on the total cost of the project was found to be 2% for every additional classroom.

It was noticed from the collected data that the concrete area for this model is so large that it might be as much as double the concrete area of any other design models available to the Ministry of Education. If the soil condition or the bearing capacity of the project force the designer to use mat foundation rather than the isolated footing, then the additional induced cost will be very high due to its very large area. The effect of this factor was found to be the greatest among all contributing factors to the model and reached 11.5% of the total construction cost of the project. It may be noticed that the execution time was the same (2 months) for all data points and that is why it did not play any role in expressing the developed statistical model.

Table 35: The Effect of Increase in one of the Parameters on the Overall Construction Cost of School Project for school design model number 11

Parameter	Amount of Increased	Percent of Increase
Length of fence	10(m)	1.1%
Type of footing	isolated to matt	11.5%
Number of classrooms	1	2%

4.3.5.2 Testing of Assumption

When a regression model is applied in practice, one cannot usually be sure in advance that the model implied is appropriate for the situation in hand. Consequently, the model considered needs to be checked for suitability and its agreement with the earlier assumptions made to develop it. As stated earlier, the following assumption were made:

i The regression function is linear

ii The distribution is normal

iii There is a constant variance over all schools construction cost.

In the following paragraphs, the above mentioned assumptions will be verified and checked.

4.3.5.2.1 Linearity Assumption

In order to test the developed model for its aptness and linearity, the residual plot was analyzed and studied. This graph was obtained by plotting the residuals against the fitted values (Figure 19). If the specified regression model is linear, the residuals will tend to scatter at random around the zero line when plotted against the fitted values. As can be seen in Figure 19 there is no pattern of systematic departure and all points are scattered normally around the zero line. This shows that the selected statistical model for all design models built by the Ministry of Education is linear and apt.

4.3.5.1.2 Normality and Constant Variance Assumptions

In order to check the normality assumption, and to be sure that a constant variance exists over all schools construction cost, the lack of normality test was performed. In this test, the standardize residual was calculated. If the distribution was normal with the constant variance and the sample size is considerably large (greater than 30 observations), the standardize residual will tend to follow the standard normal distribution figure 10. Hence, about half of the standardized residuals should be positive and half negative, About 68% of the standardized residual distribution should fall between -1 and +1 and at least 95% of the area under the normal distribution should be contained within t (0.025: n-2) and t (0.975: n-2) (i.e. $\mu \pm 2\sigma$) (Netir, Warerman and Whitemore, 1978).

In order to check if the developed model complies with the above mentioned criteria of normal distribution, the standardize residuals were calculated and listed in table 36.

Based on the analysis that was conducted on the results listed in table (36) the following was found:

t (0.025 : 37) = -.2.021

t (0.975 : 37) = +2.021

- 1 < 74% of distribution < +1

51.4% of the distribution are negative

-2.021 < 97.4% of the distribution <2.021

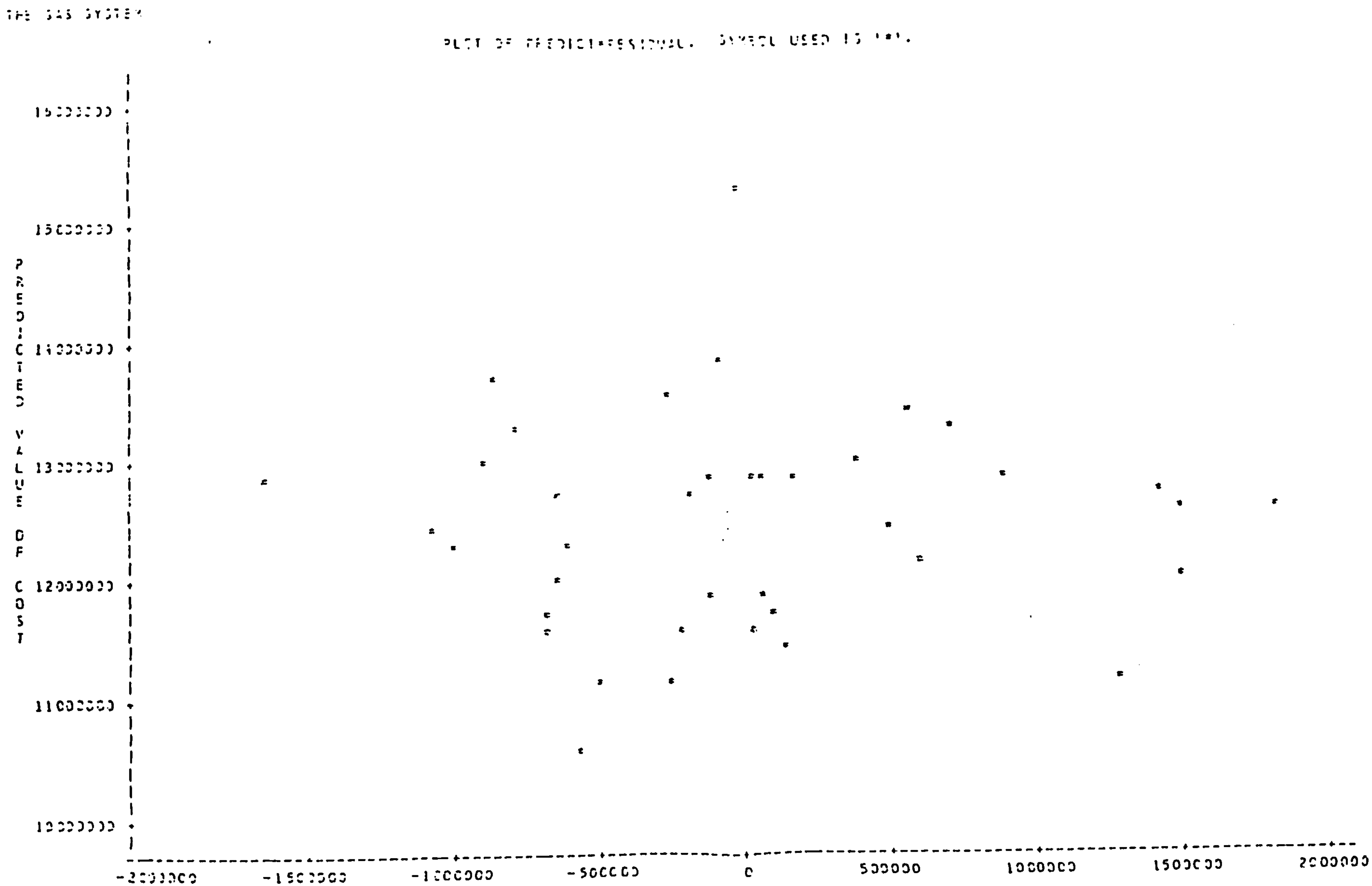

Figure 19: Residual Plot for Design Model No. 11

TABLE 36: RESIDUAL TABLE FOR REGRESSION MODEL # 11

NO	MODEL	COST	PREDICTED	RESIDUL	STANDARDISED RESIDUL
1	11	11543033	11522512	20521	0.023771187
2	11	11411937	11629922	-217985	-0.252510217
3	11	13490045	12007711	1482334	1.717111177
4	11	12356315	11078830	1277485	1.479817485
5	11	11384782	12467174	-1082392	-1.253824982
6	11	12935335	12894791	40544	0.046965499
7	11	12914745	12849053	65692	0.076096526
8	11	11421496	12060437	-638941	-0.740138682
9	11	13107765	12923039	184726	0.213983542
10	11	11339652	12336619	-996967	-1.154870076
11	11	11609838	12233688	-623850	-0.722657517
12	11	11252936	12906741	-1653805	-1.915740346
13	11	13956005	13400742	555263	0.643207471
14	11	12717476	12835278	-117802	-0.136459887
15	11	14075180	12645875	1429305	1.655683261
16	11	12869200	12368926	500274	0.579509123
17	11	12462787	13248535	-785748	-0.910197481
18	11	11018318	11713577	-695259	-0.805376521
19	11	10920579	11609387	-688808	-0.797903789
20	11	10872666	11123008	-250342	-0.28999203

Table 36: Continued

NO	MODEL	COST	PREDICTED	RESIDUL	STANDARDISED RESIDUL
20	11	10872666	11123008	-250342	-0.28999203
21	11	11961795	11903369	58426	0.067679712
22	11	13777792	12873202	904590	1.047862087
23	11	13410720	13001393	409327	0.474157624
24	11	15283084	15283084	0	0
25	11	14107256	12618805	1488451	1.724197009
26	11	12569974	12750553	-180579	-0.209179726
27	11	12071516	12729078	-657562	-0.76170894
28	11	12048217	12928842	-880625	-1.020101428
29	11	12882354	13722932	-840578	-0.973711646
30	11	13918823	13216635	702188	0.81340296
31	11	14398474	12576535	1821939	2.110503989
32	11	12758486	12143974	614512	0.711840532
33	11	9995390	10584436	-589046	-0.682341139
34	11	11810756	11686994	123762	0.143363853
35	11	10643880	11147362	-503482	-0.583225218
36	11	11684579	11796080	-111501	-0.129160913
37	11	13722367	13789082	-66715	-0.077281552
38	11	13288946	13545577	-256631	-0.297277104
39	11	11562856	11403575	159281	0.184508475

Since the above analysis complies with the different test criteria, then it can be concluded that this analysis provides no evidence of any departure form normality and the variance is constant over all school construction cost.

4.3.5.3 Validity of the Model

In order to test or to verify the validity of the model, six schools were selected randomly, excluded from the data base and withheld for checking the predictability of the developed model. The reliability of the model's predictability was examined by comparing the bid price with the model estimates (Table 37). As can be noticed, the model tends to under estimate the project cost. This might be due to the overlook of some factors that governs the construction cost of this specific design model, however, these parameters seem to have less importance than those included in the developed statistical model since the percent differences listed in table 37 are considered small.

Moreover, by considering the fact that the model is primarily designed for the purpose of preliminary estimating and, moreover, the variation in predicted and actual costs falls within the acceptable limits (-30%, + 30%), then it can be concluded that the developed model is the most representative realistic model for school design model # 11.

Table 37: A Comparison Between the Bid Price and Predicted Cost for school design model number 11

Project Name	Predicted Cost	Bid Price	Percent Difference
Al-Dudas*	11029590*	11525105	-4%
Hara	13371547	14562267	- 9%
Al-Farouq	13732448	14765769	- 7%
Al-Tahawi	1301998	12771480	+ 2%
Obada Bin Bishr	11624314	11982961	-3%
Salman Prince	15477829	13508010	-4%

* length of fence = 380 (m)
foot type = 1
number of classrooms = 20

Predicted Cost = 4194245.82 + 6779.43 (380) + 1229399.4 (1) + 151488.12 (20) = 11029540(SR)

4.3.5.4 *Predictability of the Model*

In order to compare the predictability of the developed model against the estimated cost obtained by using the current estimating system of the Ministry of Education, a sample of schools was randomly selected. As can be seen in Table 38, the estimated cost using the developed model was greatly improved. This is thought to be due to the involvement of several influencing variables in the statistical model rather than only one variable (concrete area) as in the current estimating system used by the Ministry of Education. Another reason that might be behind the improvement obtained in the estimated cost is the great importance given to the potential cost factors in any construction project, such as the site conditions, or to the special factors which distinguish school projects from any other projects, such as its level (i.e. elementary, intermediate or secondary) and the available number of technical laboratories.

Table 38: Comparison Between the Predicted Cost Using the Developed Regression Model # 11 and the Cost Estimate Prepared by the Ministry of Education

Project Name	Design Model	Bid Price (1)	Predicted Cost Using the Regression Model (2)	Estimated Cost Using the Ministry's System(3)	Percentage Difference Between 1 & 2	Percentage Difference Between 1 & 3
AL-QUDAS	11	11525105	11029590	9520000	-4%	-17%
HARA	11	14562267	13371547	11430000	-9%	-21%
AL-FAROUQ	11	14765769	13732448	9520000	-7%	-35%
AL-TAHAMI	11	12771480	1301998	11112500	+2%	-13%
OBADA BIN BASHIR	11	11982961	11624314	11112500	-3%	-7%
SALMAN PRINCE	11	13508010	15477829	11430000	-4%	-15%

4.3.5.5 Confidence Level of the Model

Confidence level can be interpreted as the probability of the estimate being correct. As being known, a good estimate implies a small length of confidence interval at high level of confidence. Thus, to measure the interval at which our estimate is correct, the upper and lower confidence limit were calculated at 95% confidence level. Moreover, the percentage of the upper and lower limits of the predicted cost with respect to its mean were calculated to have an estimate of how the predicted cost may vary within the above specified level of confidence. In table 39, the predicted costs and their statistical analysis as described above are presented.

By referring to table (39), it can be noticed that at 95% confidence level the predicted costs may vary with ± 11% only from their mean value. This low confidence interval is considered as an indication for the goodness and reliability of the estimated cost obtained from the developed statistical model.

TABLE 39 : CONFIDENCE LEVEL FOR REGRESSION MODEL # 11

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
1	11522512	244805.9	11024451	12020574	0.956774964	1.043225123
2	11629922	221700.9	11178868	12080976	0.961216077	1.038783923
3	12007711	300903.7	11395517	12619904	0.949016594	1.050983322
4	11078830	383375.2	10298848	11858813	0.929597078	1.070403012
5	12467174	214364.9	12031045	12903302	0.965017814	1.034982106
6	12894791	205759.8	12476170	13313413	0.967535651	1.032464427
7	12849053	230275.9	12380553	13317552	0.963538169	1.036461753
8	12060437	333562.6	11381799	12739076	0.943730231	1.056269852
9	12923039	198817.4	12518542	13327536	0.968699545	1.031300455
10	12336619	247666.6	11832738	12840501	0.959155665	1.040844416
11	12233688	272033.5	11680232	12787144	0.954759677	1.045240323
12	12906741	350901.0	12192828	13620654	0.944686811	1.055313189
13	13400742	381871.9	12623818	14177666	0.942023807	1.057976193
14	12835278	293661.0	12237820	13432736	0.953451885	1.046548115
15	12645875	300811.3	12033870	13257880	0.951604377	1.048395623
16	12368926	231128.1	11898692	12839160	0.961982633	1.038017367
17	13248535	331248.7	12574605	13922466	0.949131734	1.050868341
18	11713577	245388.8	11214329	12212824	0.95737869	1.042621225
19	11609387	202109.5	11198192	12020582	0.964580817	1.035419183
20	11123008	401313.5	10306529	11939486	0.926595486	1.073404424

Table 39: Continued

OBS NO.	PREDICTED COST	STD ERR. PREDICT	LOWER 95% MEAN	UPPER 95% MEAN	% OF LOWEST PRED.COST	% OF HIGHEST PRED.COST
21	11903369	327281.2	11237511	12569228	0.944061383	1.055938701
22	12873202	208701.7	12448596	13297809	0.967016287	1.032983791
23	13001393	189360.0	12616137	13386648	0.970368098	1.029631825
24	15283084	851637.9	13550414	17015754	0.886628249	1.113371751
25	12618805	300862.9	12006695	13230915	0.951492237	1.048507763
26	12750553	244128.5	12253870	13247236	0.96104616	1.03895384
27	12729078	241112.8	12238531	13219626	0.961462488	1.038537591
28	12928842	232436.4	12455946	13401737	0.963423174	1.036576748
29	13722932	463703.9	12779519	14666344	0.931252811	1.068747116
30	13216635	330191.3	12544856	13888415	0.949171707	1.050828369
31	12576535	316254.3	11933110	13219959	0.948839247	1.051160673
32	12143974	293032.1	11547796	12740152	0.950907504	1.049092496
33	10584436	361489.4	9848980	11319892	0.930515334	1.069484666
34	11686994	219228.6	11240970	12133018	0.961835866	1.038164134
35	11147362	425593.7	10281485	12013239	0.922324493	1.077675507
36	11796080	432053.8	10917060	12675101	0.925482025	1.07451806
37	13789082	536748.5	12697059	14881105	0.920805243	1.079194757
38	13545577	359269.8	12814637	14276517	0.946038474	1.053961526
39	11403575	242884.1	10909424	11897726	0.95666701	1.04333299

4.3.5.6 Fitting of Regression Line

As is known, the regression function represented by its plotted line provides a good fit to the sample data if it results in small residuals. So, to measure the degree of fitting of the developed model to the sample data, the residuals listed in Table 36 were carefully studied and analyzed. Moreover, the regression function was also plotted as seen in Figure (20). It was found that the bid price data represented by their plotted points were distributed on both sides of the fitted lines. Moreover, the dispersion was found to be small and in an average range of 5 % of the actual cost. Hence, it can be concluded that the regression function provides a good fit to the sample data.

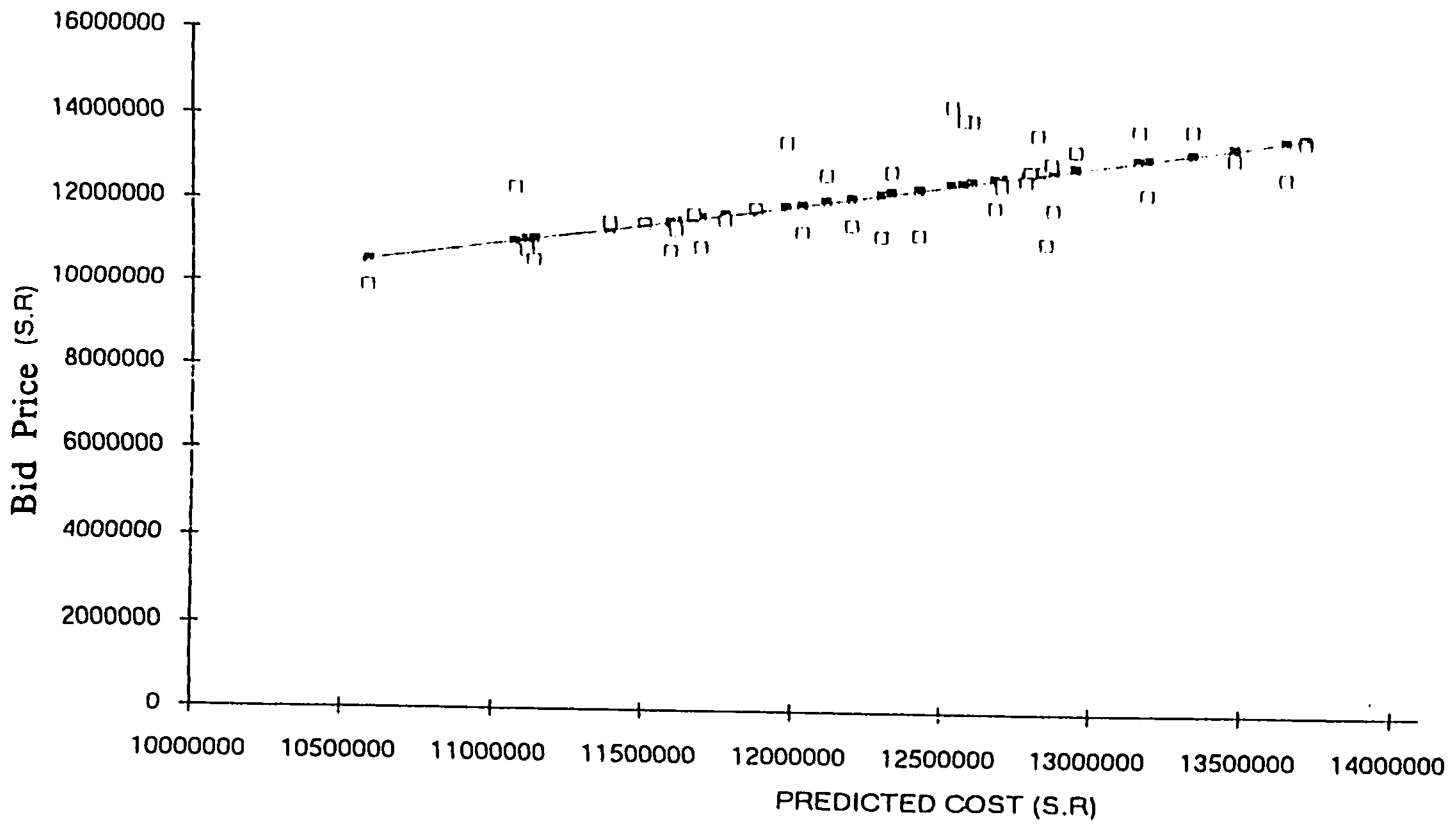

Figure 20: Regression Line for Design Model No. 11

CHAPTER FIVE

This chapter presents a summary of the study, the major findings and recommendations.

5.1 SUMMARY OF THE STUDY

The Ministry of Education builds schools all over the Kingdom. The building process involves requesting a budget from the Ministry of Finance and Economics, which is based on an estimate for the cost of building.

This study attempts to identify and discuss the current cost estimating technique used by the Ministry of Education for building their school projects through personal interviews. Upon analyzing the estimating technique, major problems and disadvantages were found. Accordingly, a trial was made to develop a new statistic- based estimating system that will be able to calculate the construction cost of school building. Models were developed from actual data obtained from 170 school project records. The proposed new system offers an easy way for preparing budget estimates for school projects, and also provides a solution for most problems encountered in the currently used system by the Ministry of Education.

5.2 SUMMARY OF MAJOR FINDINGS

1. The Ministry of Education uses simple and purely manual cost estimating system which is based on the concept of floor area unit estimate.

2. The Ministry of Education has low level of satisfaction for the currently used estimating system

3. It was possible to develop five regression models for estimating the construction cost of school projects. One model is general and can be used for design models number 2,3,5 and 11 and the other four models were developed for each design model separately.

4. In the general model, the concrete area, footing type, number of technical laboratories, quantity of back filling and excavated materials were found to have a significant impact on the cost of school building projects. These parameters explained 90% of the variation in the construction cost of school projects.

5. The general model is reliable and the accuracy of its estimate ranges between 2.5% to 18.5%.

6. In the statistical model developed for design model number 2, the soil improvement, number of technical laboratories, quantity of excavated and back filling materials, length of fence and concrete area were found to have a significant impact on the cost of the school building project. These parameters explained about 59% of the variation in the construction cost of school project built as per this design model.

7. The developed statistical model for design model number 2 is reliable and the accuracy of its estimate ranges between 9% to 24%.

8. In the statistical model developed for design model number 3, the excavated material, back filling materials and level of school were found to have a significant impact on the cost of school building projects. These parameters explained 87% of the variation in the construction cost of school project built as per this design model.

9. The developed statistical model for design model number 3 is reliable and the accuracy of its estimate ranges between 2% to 19.5%.

10. In the statistical model developed for design model number 5, the number of class rooms, the length of retaining walls and the level of school were found to have a significant impact on the cost of school building projects. These parameters explained 66% of the variation in the construction cost of school projects built as per this design.

11. The developed statistical model for design model number 5 is reliable and the accuracy of its estimate ranges between 1% to 8%.

12. In the statistical model developed for design model number 11, the length of fence, the foundation type and the number of class rooms were found to have a significant impact on the cost of the school building projects. These parameters explained 52% of the variation is the construction cost of school projects built as per this design model.

13. The developed statistical model for design model number 11 is reliable and the accuracy of its estimate ranges between 2% and 9%.

14. The construction cost of school projects was found to be heavily dependent on some factors that are not included in the design such

as quantity of excavated materials, quantity of backfilling material, and soil improvement techniques.

15. It was found that the province in which the school is constructed has a great impact on the nature of independent variables governing the total cost of the project as in Model numbers 5 and 11.

16. The over-design problem encountered in some design models was the basic reason for underevaluating the importance of some variables in expressing the regression equation.

5.3 RECOMMENDATIONS

5.3.1 General Recommendations

1. The Ministry of Education is recommended to use the developed regression models for preparing budget estimates for school building projects.

5.3.2 Recommendations for Future Studies

1. Since the refinement process of almost any research work is somewhat endless, it is recommended to include more parameters in the development models such as cost index, so that their predictability can be improved and projected to the future.

2. It is recommended for the same study to be implemented on the private and girls schools projects, so that buduget estimating systems for different types of schools in the Kingdom of Saudi Arabia is developed.

APPENDICES

APPENDIX - A

Table 1: Information Needed on Parameters for Every School

School Name: ______________________

Level of School: ______________________

Construction Year ______________________

School Design Model ______________________

Total Enclosed Area (m^2)	
Availability of Technical Lab.	[ ] Yes [ ] No
Number of Technical Laboratory	
Total Number of Classrooms	
Construction Time (Months)	
Total Cost (SR)	
Length of Retaining Walls (m)	
Is There any Soil Improvement	[ ] Yes [ ] No
Length of Fence (m)	
Bearing Capacity of Soil	
Footing Types	[ ] Isolated [ ] Matt [ ] Strip
Amount of Excavated Materials (m^3)	
Amount of Backfilling Materials (m^3)	

Table 2: Information Needed on Parameters for All Schools

SCHOOL NO	SCHOOL NAME	CONSTRUTIION TIME	TOTAL ENCLOSED AREA	NO. OF CLASS ROOMS	LENGTH OF FENCE	SOIL IMPROVE MENT	LENGTH OF RETAINING WALLS	AVAILABILITY OF TECH. LABS	TOTAL COST	FOOTIG TYPE	LEVEL OF SCHOOL	NO. OF TECH. LABS.	AMOUNT OF EXCAVATED MATERIAL	BEARING CAPACITY OF SOIL	AMOUNT OF BACK FILLING MATERIAL

QUESTIONNAIRE

How do you prepare the budget estimate for building new school projects (steps and process)?

When it was developed	
Who develop it	
What are its Input Data	1- ______ 7 ______ 2- ______ 8 ______ 3- ______ 9 ______ 4- ______ 10 ______ 5- ______ 11 ______ 6- ______ 12 ______
For how long it has been in use	
Type of system	[ ] Manual [ ] Computerized
Who is performing the Estimate	[ ] Engineers [ ] Non Engrs
The system gives detailed estimate The system gives lumpsum estiamte	[ ] [ ]
In case the system gives detailed estimate, what are the basic division	1- ______ 6 ______ 2- ______ 7 ______ 3- ______ 8 ______ 4- ______ 9 ______ 5- ______ 10 ______
Is the system applicable for all type of schools:	[ ] Yes [ ] No
In case the system is not applicable for all types of schools, what are the different systems you have	1 - 2 - 3 -

Who approves the estimate and what are the criteria of approving	
If the system is manual, is there any standard format.	[ ] Yes [ ] No
What is the organizational chart for the estimating department.	
What are the steps followed in preparing the budget estimate	1- 2- 3- 4- 5- 6- 7- 8- 9- 10-
What are the steps followed to get the approval on the budget from the Ministry of Finance	1- 2- 3- 4- 5-

What is the over all satisfaction level acquired form the system	High [ ] Moderate [ ] No satisfaction [ ]
What are the areas of problems that is felt by the end users (Disadvantages)	1- 2- 3- 4- 5- 6- 7- 8- 9- 10-
What is the approximate time to make the estiamte	
What is the maximum allowable time to make the estiamte, if a new system will be developed.	
Advantages	1- 2- 3- 4- 5- 6- 7- 8- 9- 10-

APPENDIX - B

CORRELATION MATRIX FOR GENERAL MODEL

VARIABLE	NO. OF CLASS ROOMS	BID PRICE	AREA OF SITE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY
NO. OF CLASS ROOMS	1.00000	0.43849	0.23439	0.22578	0.64177	0.59870	-0.07440
BID PRICE	0.43849	1.00000	0.76702	0.72190	0.41895	0.14660	0.05071
AREA OF SITE	0.23539	0.76702	1.00000	0.94082	0.17651	0.19730	0.07971
LEGTH OF FENCE	0.22578	0.72190	0.94082	1.00000	0.14483	0.20306	0.09067
EXCUTION TIME	0.64177	0.41895	0.17651	0.14483	1.00000	0.48300	-0.06229
FOOTING TYPE	0.05980	0.14660	0.19730	0.20306	0.04830	1.00000	-0.39174
BEARING CAPACITY	-0.07440	0.05071	0.07971	0.09067	0.06229	0.39174	1.00000
SOIL IMPROVEMENT	0.07374	0.04888	0.05158	0.06392	0.10080	0.47917	0.42336

CORRELATION MATRIX FOR GENERAL MODEL

VARIABLE	NO. OF CLASS ROOMS	BID PRICE	AREA OF SITE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY
EXCAVATION	0.19616	0.27736	0.23204	0.26186	0.11412	-0.07456	0.14830
FILL	0.21433	0.52917	0.54700	0.49618	0.18728	0.13488	0.10263
LEVEL	-0.08297	0.41859	0.37222	0.32332	0.18486	0.03992	0.07539
LAB NO	-0.08297	0.41859	0.37216	0.32332	0.18486	0.03992	0.07539
CONCRETE AREA	0.61318	0.91794	0.64036	0.60702	0.53304	0.05144	0.05514
RW	0.17881	0.36966	0.29846	0.31858	0.16785	0.01716	0.19222

CORRELATION MATRIX FOR GENERAL MODEL

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
NO. OF CLASS ROOM	0.07374	0.19616	0.21433	-0.08297	-0.08297	0.61318	0.17881
BID PRICE	0.04888	0.27736	0.52917	0.41859	0.41859	0.91794	0.36966
AREA OF SITE	0.05158	0.23204	0.54700	0.37216	0.37216	0.64036	0.29846
LEGTH OF FENCE	0.06392	0.26186	0.49618	0.32332	0.32332	0.60702	0.31858
EXCUTION TIME	0.10080	0.11412	0.18728	0.18486	0.18486	0.53304	0.16785
FOOTING TYPE	0.47917	-0.07456	0.13488	0.03992	0.03992	0.05144	0.01716
BEARING CAPACITY	-0.42336	0.14830	0.10263	0.07539	0.07539	0.05514	0.19222
SOIL IMPROVEMENT	1.00000	-0.01983	-0.02652	0.03514	0.03514	0.00269	-0.07180

CORRELATION MATRIX FOR GENERAL MODEL

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
EXCAVATION	-0.01983	1.00000	0.12415	0.16072	0.16072	0.22832	0.54926
FILL	-0.02652	0.12415	1.00000	0.09645	0.09645	0.45162	0.66450
LEVEL	0.03514	0.16072	0.09645	1.00000	1.00000	0.27272	0.08880
LAB NO	0.03514	0.16072	0.09645	1.00000	1.00000	0.27272	0.08880
CONCRETE AREA	0.00269	0.22832	0.45162	0.27272	0.27272	1.00000	0.27977
RW	-0.07180	0.54926	0.66450	0.08880	0.08880	0.27977	1.00000

CORRELATION MATRIX FOR REGRESSION MODEL NO. 2

VARIABLE	NO.OF CLASS ROOMS	BID PRICE	AREA OF SITE	LEGHT OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY
NO. OF CLASS ROOMS	1.00000	0.03009	-0.23248	-0.12704	0.98113	0.08674	-0.10107
BID PRICE	0.03093	1.00000	0.48251	0.46466	0.37298	0.41248	-0.10535
AREA OF SITE	-0.23248	0.48251	1.00000	0.90860	-0.20193	0.19999	0.07819
LEGTH OF FENCE	-0.12704	0.46466	0.90860	1.00000	-0.10443	0.23996	0.08878
EXCUTION TIME	0.98113	0.37298	-0.20193	-0.10443	1.00000	0.07865	-0.07651
FOOTING TYPE	0.08674	0.41248	0.19989	0.23996	0.07865	1.00000	-0.54016
BEARING CAPACITY	-0.10107	-0.10535	0.07819	0.08878	-0.07651	-0.54016	1.00000
SOIL IMPROVEMENT	0.08934	0.38339	0.18933	0.18206	0.14274	0.64343	-0.40113

CORRELATION MATRIX FOR REGRESSION MODEL NO. 2

VARIABLE	NO.OF CLASS ROOMS	BID PRICE	AREA OF SITE	LEGHT OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY
EXCAVATION	0.06990	0.32795	0.17891	0.19418	0.07391	-0.09974	0.06214
FILL	-0.00626	0.31463	0.36234	0.32220	-0.00725	0.02047	0.07828
LEVEL	-0.29578	0.22962	0.21649	0.15921	-0.21602	-0.04997	0.19834
LAB. NO	-0.29578	0.22962	0.21649	0.15921	-0.21602	-0.04997	0.19834
CONCRETE AREA	0.97849	0.38488	-0.19442	-0.09764	0.09991	0.07769	-0.06950
RW	0.12343	0.48021	0.25935	0.28216	0.14406	-0.05954	0.13854

CORRELATION MATRIX FOR REGRESSION MODEL NO. 2

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
NO.OF CLASS ROOMS	0.08934	0.06990	-0.00626	-0.29478	-0.29578	0.97849	0.12343
BID PRICE	0.38339	0.32795	0.31463	0.22962	0.22962	0.38488	0.48021
AREA OF SITE	0.18933	0.17891	0.36234	0.21649	0.21649	-0.19442	0.25935
LEGTH OF FENCE	0.18206	0.19418	0.32220	0.15921	0.15921	-0.09764	0.28216
EXCUTION TIME	0.14274	0.07391	-0.00725	-0.21602	-0.21602	0.09991	0.14406
FOOTING TYPE	0.64343	-0.99740	0.02047	-0.04997	-0.04997	0.07769	-0.05954
BEARING CAPACITY	-0.40113	0.06214	0.07828	0.19834	0.19834	-0.06950	0.13854
SOIL IMPROVEMENT	1.00000	-0.05650	0.04316	-0.02071	-0.02071	0.14055	-0.02077

CORRELATION MATRIX FOR REGRESSION MODEL NO. 2

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
EXCAVATION	-0.05650	1.00000	0.01247	0.15592	0.15592	0.08242	0.65986
FILL	0.04316	0.01247	1.00000	0.03522	0.03522	-0.00497	0.57143
LEVEL	-0.02071	0.15592	0.03522	1.00000	1.00000	-0.17698	0.13924
LAB NO	-0.02071	0.15592	0.03522	1.00000	1.00000	-0.17698	0.13924
CONCRETE AREA	0.14055	0.08242	-0.00497	-0.17698	-0.17698	1.00000	0.15238
RW	-0.02077	0.65986	0.57143	0.13924	0.13924	0.15238	1.00000

CORRELATION MATRIX FOR REGRESSION MODEL NO. 3

VARIABLE	NO.OF CLASS ROOMS	BID PRICE	LENGTH OF FENCE	EXCUTION TIME	BEARING CAPACITY
NO.OF CLASS ROOMS	1.00000	0.19419	0.13065	0.55922	-0.15719
BID PRICE	0.19419	1.00000	0.05318	0.73592	-0.30655
LENGTH OF FENCE	0.13065	0.05318	1.00000	0.06346	0.04738
EXCUTION TIME	0.55922	0.73592	0.06346	1.00000	-0.24564
BEARING CAPACITY	-0.15719	-0.30650	0.04738	-0.24564	1.00000
SOIL IMPROVEMENT	0.34176	0.01746	-0.12684	0.17609	-0.23250
EXCAVATION	0.44185	0.38810	0.39134	0.22076	-0.13390

CORRELATION MATRIX FOR REGRESSION MODEL NO. 3

VARIABLE	NO.OF CLASS ROOMS	BID PRICE	LENGTH OF FENCE	EXCUTION TIME	BEARING CAPACITY
FILL	0.11820	0.33774	0.03212	0.26609	0.08852
LEVEL	-0.05031	0.84417	-0.06229	0.62413	-0.27058
LAB NO	-0.05031	0.84417	0.06289	0.62413	-0.27058
CONCRETE AREA	0.50861	0.74641	0.06313	0.94446	-0.23697
RW	0.08769	0.39695	0.24012	0.16716	0.22620

CORRELATION MATRIX FOR REGRESSION MODEL NO. 3

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
NO.OF CLASS ROOMS	0.34176	0.44185	0.11820	-0.05031	-0.05031	0.50861	0.08769
BID PRICE	0.01746	0.38810	0.33774	0.84417	0.84417	0.74641	0.39695
LENGTH OF FENCE	0.12684	0.39134	0.03212	-0.06289	-0.06289	0.06313	0.24012
EXCUTION TIME	0.17609	0.22076	0.26609	0.62413	0.62413	0.94446	0.16716
BEARING CAPACITY	-0.23250	-0.13390	0.08852	-0.27058	-0.27058	-0.23697	0.22620
SOIL IMPROVEMENT	1.00000	0.28824	0.03920	-0.18809	-0.18809	0.14799	-0.09903
EXCAVATION	0.28824	1.00000	0.01747	0.17856	0.17856	0.27945	0.41394

CORRELATION MATRIX FOR REGRESSION MODEL NO. 3

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
FILL	0.03920	0.01747	1.00000	0.02061	0.02061	0.25296	0.66128
LEVEL	-0.18809	0.17856	0.02061	1.00000	1.00000	0.66626	0.15301
LAB NO	-0.18809	0.17856	0.02061	1.00000	1.00000	0.66626	0.15301
CONCRETE AREA	0.14799	0.27945	0.25296	0.66626	0.66626	1.00000	0.18112
RW	-0.09903	0.41394	0.66128	0.15301	0.15301	0.18112	1.00000

CORRELATION MATRIX FOR REGRESSION MODEL NO. 5

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
NO.OF CLASS ROOMS	-0.32684	0.23888	-0.14575	-0.17303	-0.17303	0.88869	0.21312
BID PRICE	-0.30109	0.50486	0.20545	0.24800	0.24800	0.53797	0.69892
AREA OF SITE	-0.22510	0.02071	0.49525	0.71027	0.71027	0.43615	0.13022
LENGTH OF FENCE	-0.17940	0.04185	0.33161	0.18026	0.18026	0.25445	0.31238
EXCUTION TIME	-0.08193	0.11269	-0.26001	-0.08193	-0.08193	0.68113	-0.09855
BEARING CAPACITY	-0.46953	0.34837	-0.12977	-0.28172	-0.28172	0.02831	0.36256
SOIL IMPROVEMENT	1.00000	-0.22697	0.12470	-0.11111	-0.11111	-0.36851	-0.08517
EXCAVATION	-0.22697	1.00000	0.06185	-0.13194	-0.13194	0.16937	0.69144

CORRELATION MATRIX FOR REGRESSION MODEL NO. 5

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
FILL	0.12470	0.06185	1.00000	0.67003	0.67003	0.17316	0.44397
LEVEL	-0.11111	-0.13194	0.67003	1.00000	1.00000	0.30151	0.01820
LAB NO	-0.11111	-0.13194	0.67003	1.00000	1.00000	0.30151	0.01820
CONCRETE AREA	-0.36851	0.16937	0.17316	0.30151	0.30151	1.00000	0.21485
RW	-0.08517	0.69144	0.44397	0.01820	0.01820	0.21485	1.00000

CORRELATION MATRIX FOR REGRESSION MODEL NO. 5

VARIABLE	NO.OF CLASS ROOMS	BID PRICE	AREA OF SITE	LENGTH OF FENCE	EXCUTION TIME	BEARING CAPACITY
NO.OF CLASS ROOMS	1.00000	0.43556	0.10642	0.17551	0.74329	0.16574
BID PRICE	0.43556	1.00000	0.42789	0.35341	0.25537	0.36908
AREA OF SITE	0.10642	0.42789	1.00000	0.48180	0.19972	-0.14189
LENGTH OF FENCE	0.17551	0.35341	0.48180	1.00000	0.00405	-0.10385
EXCUTION TIME	0.74329	0.25537	0.19972	0.04046	1.00000	0.06925
BEARING CAPACITY	0.16574	0.36908	-0.14189	-0.10385	0.06925	1.00000
SOIL IMPROVEMENT	-0.32684	-0.30109	-0.22510	-0.17940	-0.08193	-0.46953
EXCAVATION	0.23888	0.50486	0.02071	0.04185	0.11269	0.34837

CORRELATION MATRIX FOR REGRESSION MODEL NO. 5

VARIABLE	NO.OF CLASS ROOMS	BID PRICE	AREA OF SITE	LENGTH OF FENCE	EXCUTION TIME	BEARING CAPACITY
FILL	-0.14575	0.20545	0.49525	0.33161	-0.26001	-0.12977
LEVEL	-0.17303	0.24800	0.71027	0.18026	-0.08193	-0.28172
LAB NO	-0.17303	0.24800	0.71027	0.18026	-0.08193	-0.28172
CONCRETE AREA	0.88691	0.53797	0.43615	0.25445	0.68113	0.02831
RW	0.21312	0.69892	0.13022	0.31238	-0.09855	0.36256

CORRELATION MATRIX FOR REGRESSION MODEL NO. 11

VARIABLE	NO.OF CLASS ROOMS	BID PRICE	AREA OF SITE	LENGTH OF FENCE	FOOTING TYPE	BEARING CAPACITY
NO.OF CLASS ROOM S	1.00000	0.50212	0.35372	0.36916	-0.12969	0.13430
BID PRICE	0.50212	1.00000	0.61721	0.59158	0.38657	-0.04212
AREA OF SITE	0.35372	0.61721	1.00000	0.92753	0.34493	-0.00591
LENGTH OF FENCE	0.36916	0.59158	0.92753	1.00000	0.36892	-0.08283
FOOTING TYPE	-0.12969	0.38657	0.34493	0.36892	1.00000	-0.34785
BEARING CAPACITY	0.13430	-0.04212	-0.00591	-0.08283	-0.34785	1.00000
SOIL IMPROVEMENT	0.10618	0.17207	0.15096	0.21984	0.47986	-0.61091
EXCAVATION	0.18517	0.17289	-0.06436	0.00069	-0.10631	0.29490

CORRELATION MATRIX FOR REGRESSION MODEL NO. 11

VARIABLE	NO.OF CLASS ROOMS	BID PRICE	AREA OF SITE	LENGTH OF FENCE	FOOTING TYPE	BEARING CAPACITY
FILL	0.05264	0.34409	0.38732	0.35865	0.20130	0.14649
LEVEL	0.58408	0.49925	0.29426	0.28491	0.19145	-0.16327
LAB NO	0.58408	0.49925	0.29426	0.28491	0.19145	-0.16327
CONCRETE AREA	0.72219	0.42176	0.18511	0.22890	0.05484	0.04092
RW	0.04841	0.13448	0.09444	0.12575	0.07720	0.21749

CORRELATION MATRIX FOR REGRESSION MODEL NO. 11

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
NO.OF CLASS ROOMS	0.10618	0.18517	0.05264	0.58408	0.58408	0.77219	0.04841
BID PRICE	0.17207	0.17289	0.34409	0.49925	0.49925	0.42176	0.13448
AREA OF SITE	0.15096	-0.06436	0.38732	0.29426	0.29426	0.18511	0.09444
LENGTH OF FENCE	0.21984	0.00069	0.35865	0.28491	0.28491	0.23289	0.12575
FOOTING TYPE	0.47986	-0.10631	0.20130	0.19145	0.19145	0.05484	0.07720
BEARING CAPACITY	-0.61092	0.29490	0.14649	-0.16333	-0.16327	0.04092	0.21749
SOIL IMPROVEMENT	1.00000	-0.04906	-0.12068	0.39898	0.39898	0.11429	-0.13846
EXCAVATION	-0.04906	1.00000	0.02895	0.02059	0.02059	0.12935	0.41003

CORRELATION MATRIX FOR REGRESSION MODEL NO. 11

VARIABLE	SOIL IMPROVEMENT	EXCAVATION	FILL	LEVEL	LAB NO	CONCRETE AREA	RW
FILL	-0.12068	0.02895	1.00000	-0.13355	-0.13355	0.01062	0.77724
LEVEL	0.39898	0.02059	0.13355	1.00000	1.00000	0.39898	-0.23125
LAB NO	0.39898	0.02059	0.13355	1.00000	1.00000	0.39898	-0.23125
CONCRETE AREA	0.11429	0.12935	0.01062	0.39898	0.39898	1.00000	0.09486
RW	-0.13846	0.41003	0.77238	-0.23150	-0.23150	0.09486	1.00000

APPENDIX - C

Table (1): Number of Observations Associated With Each Design Modle

Modle Number	Number of Observation
2	83
3	36
4	0
5	24
6	2
10	2
11	45

Total Observation = 192

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
1	3	22	5553587	302	24	1	2.0	0	600	2950
2	2	28	6904694	360	24	1	2.0	0	0	19200
3	2	18	5436815	320	21	1	2.0	0	70	1250
4	3	12	4642305	260	18	1	2.0	0	600	3000
5	2	18	5733160	350	21	1	2.0	0	700	4700
6	3	22	6374790	330	24	1	2.0	0	650	12900
7	3	12	6901734	200	24	1	2.0	0	0	2800
9	5	21	6559178	150	24	1	1.5	0	650	2700
10	2	19	5073788	180	21	1	2.0	0	400	250

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3172
2	1	0	0	21	161	96	10	0	0	0	0	0	288	0	3670
3	1	101	0	0	0	0	0	0	0	0	0	0	101	0	2553
4	1	41	0	0	0	0	0	0	0	0	0	0	41	0	2221
5	1	15	0	0	0	0	0	0	0	0	0	0	15	0	2553
6	1	0	15	55	15	75	20	0	0	0	0	0	180	0	3172
7	3	0	0	0	0	0	25	28	0	0	0	0	53	4	3220
9	2	30	0	0	0	0	0	0	0	0	0	0	30	2	3453
10	1	35	0	0	0	0	0	0	0	0	0	0	35	0	2536

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
11	5	28	7176342	236	24	1	2.0	0	2450	2100
12	3	22	4961254	370	21	1	2.0	0	3550	1100
13	11	25	11543033	350	24	1	2.0	0	50	9500
14	11	25	11411937	390	24	1	2.0	0	550	2900
15	11	25	13490045	450	24	1	1.0	0	50	5650
16	11	20	12356315	375	24	1	1.0	0	0	16900
18	6	24	24544157	687	24	1	1.5	0	400	18000
19	3	12	6791200	340	24	1	2.0	0	1250	61000
20	2	18	6392235	318	21	1	2.0	0	4600	6150

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
11	1	0	25	10	31	0	10	37	20	20	22	0	175	0	3453
12	2	0	0	0	0	0	0	0	0	0	0	0	0	2	3220
13	2	0	200	0	0	0	0	0	0	0	0	0	200	2	6350
14	2	0	0	0	0	0	0	0	0	0	0	0	0	2	6350
15	1	0	0	0	0	0	0	0	0	0	0	0	0	0	6350
16	1	0	175	0	0	0	0	0	0	0	0	0	175	0	5600
18	3	0	0	170	50	0	0	0	0	0	0	0	220	4	8801
19	3	66	52	0	0	0	0	0	0	0	0	0	118	4	3220
20	2	30	35	20	130	20	10	10	0	0	0	0	255	2	2603

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
21	2	28	5874050	308	24	1	2.0	0	1450	3250
22	11	30	11384782	384	24	1	2.0	0	2250	6600
23	2	18	5868298	250	21	1	2.0	0	160	5550
24	2	28	5329678	246	24	1	2.0	0	100	2000
25	2	28	8742473	370	24	1	1.0	0	20000	2800
26	3	22	7883848	372	24	1	2.0	0	7500	8750
28	3	12	7229710	233	24	1	2.0	0	1500	5800
29	5	19	5271682	181	21	1	1.5	0	15	900
30	3	12	4863248	350	18	1	2.0	0	1550	600

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
21	1	50	65	35	0	0	0	0	0	0	0	0	150	0	3670
22	1	85	42	26	4	0	0	0	0	0	0	0	157	0	6350
23	2	30	35	6	0	0	0	0	0	0	0	0	71	2	2603
24	1	40	10	30	20	0	0	0	0	0	0	0	100	0	3670
25	2	41	57	50	0	40	54	64	0	0	40	82	428	2	3720
26	2	93	69	63	41	0	41	15	13	5	10	170	520	2	3220
28	3	15	30	45	0	0	0	0	0	0	0	0	90	4	3220
29	1	44	0	0	0	0	0	0	0	0	0	0	44	0	2536
30	1	118	0	0	0	0	0	0	0	0	0	0	118	0	2221

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
31	2	28	7268715	330	24	1	2.0	0	470	4100
33	11	30	12935335	450	24	1	2.0	0	2000	9700
34	11	30	12914745	450	24	1	2.0	0	400	10200
35	11	30	11421496	340	24	1	1.5	0	50	2650
36	11	30	13107765	450	24	1	2.0	0	4100	7500
37	2	28	5833170	340	24	1	2.0	0	3600	850
38	11	30	11339652	370	24	1	2.0	0	4600	100
39	2	18	5950785	320	21	1	1.5	0	850	1950
40	2	18	6337282	420	21	1	2.0	0	4000	3150

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
31	1	122	10	0	0	0	0	0	0	0	0	0	132	0	3670
33	2	45	22	0	0	0	0	0	0	0	0	0	67	2	6350
34	2	0	256	0	0	0	0	0	0	0	0	0	256	2	6350
35	3	0	0	0	0	0	0	0	0	0	0	0	0	4	6350
36	2	86	32	0	0	0	0	0	0	0	0	0	118	2	6350
37	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
38	1	0	0	0	0	0	0	0	0	0	0	0	0	0	6350
39	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2603
40	2	77	0	0	0	0	0	0	0	0	0	0	77	2	2603

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
41	3	12	7007962	340	24	1	2.0	0	620	2000
42	3	12	6941803	300	24	1	2.0	0	1750	1650
43	2	28	6819579	268	24	1	1.5	0	400	2800
44	3	20	6617453	280	24	1	2.0	0	2000	700
45	3	20	7538941	340	24	1	1.5	0	2000	2300
46	2	28	7696680	260	24	1	1.5	0	0	9700
47	2	18	5955973	290	21	1	1.5	0	120	7300
49	2	18	5428406	294	21	1	1.5	0	60	4700
50	3	20	7698578	350	24	1	1.5	0	5900	720

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
41	3	60	0	0	0	0	0	0	0	0	0	0	60	4	3220
42	3	20	0	0	0	0	0	0	0	0	0	0	20	4	3220
43	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
44	3	0	0	0	0	0	0	0	0	0	0	0	0	4	3220
45	3	0	0	0	0	0	0	0	0	0	0	0	0	4	3220
46	2	0	0	68	136	0	0	0	0	0	0	0	204	2	3720
47	2	0	16	62	0	0	0	0	0	0	0	0	78	2	2603
49	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2553
50	3	40	0	0	0	0	0	0	0	0	0	0	40	4	3220

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
51	3	20	7566002	340	24	1	2.0	0	2650	3100
52	2	28	6395654	330	24	1	2.0	0	2300	1350
53	2	28	6594874	330	24	1	2.0	0	9800	350
54	11	30	11609838	365	24	1	1.0	1	850	2850
55	3	22	5076755	260	24	1	1.5	1	2300	2600
56	2	28	5862659	310	24	1	1.0	1	50	3600
57	2	28	6134082	340	24	1	0.7	1	1350	2400
58	6	24	12327863	425	24	1	2.0	0	400	9050
59	2	28	5985308	340	24	1	1.0	0	4800	1650

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
51	3	190	0	0	0	0	0	0	0	0	0	0	190	4	3220
52	1	57	25	0	0	0	0	0	0	0	0	0	82	0	3670
53	1	104	42	0	0	0	0	0	0	0	0	0	146	0	3670
54	3	0	0	0	0	0	0	0	0	0	0	0	0	0	6350
55	1	0	0	0	0	0	0	0	0	0	0	0	0	4	3172
56	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
57	1	30	30	0	0	0	0	0	0	0	0	0	60	0	3670
58	1	0	0	0	0	0	0	0	0	0	0	0	0	0	7459
59	1	84	0	0	0	0	0	0	0	0	0	0	84	0	3670

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
61	11	30	11252936	490	24	1	1.0	1	950	1100
62	2	18	4453455	320	21	1	1.5	0	1500	8500
63	2	18	4343296	220	21	1	2.0	0	1000	1800
66	2	28	5464115	152	24	1	1.5	0	120	1250
69	3	20	8307439	289	24	1	1.5	1	6350	5600
70	2	26	6832657	202	24	1	2.0	0	1800	1150
71	5	28	6100363	208	24	1	2.0	0	1200	1600
72	11	30	13956005	530	24	1	2.0	0	3200	10300
75	11	30	12717476	390	24	1	2.0	0	11500	7100

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
61	3	0	0	0	0	0	0	0	0	0	0	0	0	4	6350
62	1	10	23	105	0	30	0	0	0	0	0	0	168	0	2553
63	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2553
66	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
69	3	0	18	0	32	0	10	10	0	20	16	40	146	4	3220
70	3	0	0	0	0	0	0	0	0	0	0	0	0	4	3720
71	1	74	51	0	0	0	0	0	0	0	0	0	125	0	3453
72	3	0	0	0	0	0	0	0	0	0	0	0	0	4	6350
75	3	120	58	132	23	0	0	0	0	0	0	0	333	4	6350

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
76	11	30	14075180	388	24	1	1.5	1	10500	150
77	2	20	9125940	400	24	1	2.0	1	2800	4400
79	11	30	12869200	380	24	1	2.0	0	2000	3350
80	5	19	5890731	200	21	1	2.0	0	3550	3100
81	3	12	74402296	290	24	1	2.0	0	700	8500
82	3	20	7935874	300	24	1	2.0	0	3000	2800
85	2	18	5895825	320	21	1	2.0	0	1300	2250
86	11	30	12462787	470	24	1	2.0	0	0	24800
87	2	28	7655616	370	24	1	2.0	0	7700	22000

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
76	3	35	105	0	0	0	0	0	0	0	0	0	140	4	6350
77	3	102	68	0	0	0	0	0	0	0	0	0	170	4	3720
79	2	0	15	0	0	0	0	0	0	0	0	0	15	2	6350
80	1	20	0	20	20	20	10	20	10	0	10	70	200	0	2536
81	3	0	75	46	38	100	0	0	0	0	0	0	259	4	3220
82	3	8	0	0	0	0	0	0	0	0	0	0	8	4	3220
85	1	0	17	0	0	0	0	0	0	0	0	0	17	0	2553
86	1	101	83	120	0	0	0	0	0	0	0	0	304	0	6350
87	2	0	0	0	147	0	165	0	0	0	0	0	312	2	3720

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
88	2	18	5660981	360	21	1	1.5	0	7250	11900
90	11	25	11018318	370	24	1	2.0	0	7150	1220
91	5	19	4134785	180	24	1	2.0	0	100	50
92	11	25	10920579	355	24	1	1.0	0	3900	5800
93	11	20	10872666	360	24	1	2.0	0	8500	8700
94	11	25	11961795	370	24	1	2.0	0	0	22600
95	11	30	13777792	450	24	1	2.0	0	3500	6100
96	11	30	13410720	450	24	1	2.0	0	4750	10200
97	5	28	4825280	168	24	1	1.5	0	550	200

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
88	1	55	36	49	20	23	43	65	34	0	0	0	325	0	2553
90	1	80	80	0	0	0	0	0	0	0	0	0	160	0	6350
91	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2536
92	1	0	220	50	0	0	0	0	0	0	0	0	270	0	6350
93	1	0	0	95	0	0	0	0	0	0	0	0	95	0	5600
94	1	0	20	42	131	45	0	0	0	80	0	0	318	0	6350
95	2	0	57	0	0	0	0	0	0	0	0	0	57	2	6350
96	2	60	30	0	0	0	0	0	0	0	0	0	90	2	6350
97	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3453

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
98	2	28	5758779	285	24	1	2.0	0	100	2500
99	5	28	5093569	250	24	1	1.0	0	330	300
101	2	28	5718324	195	24	1	1.0	0	3900	100
102	5	28	4780251	150	24	1	1.0	0	660	0
103	2	18	5723083	410	21	1	1.5	0	2250	1300
104	5	21	6299232	300	21	1	1.0	0	0	5500
105	5	19	4434368	141	21	1	1.0	1	0	2300
106	11	25	15283084	533	24	3	1.0	1	1200	21000
107	11	30	14107256	370	24	1	2.0	0	10600	3600

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
98	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
99	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3453
101	1	22	0	0	0	0	0	0	0	0	0	0	22	0	3670
102	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3453
103	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2603
104	2	140	0	0	0	0	0	0	0	0	0	0	140	2	3453
105	1	0	94	0	0	0	0	0	0	0	0	0	94	0	2536
106	3	0	0	216	0	0	0	0	0	0	0	0	216	4	6350
107	3	0	60	0	0	0	0	0	0	0	0	0	60	4	6350

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
108	11	30	12569974	450	24	1	1.5	0	1350	3800
109	11	30	12071516	390	24	1	2.0	0	3000	16400
110	11	30	12048217	440	24	1	2.0	0	7700	4400
111	11	30	12882354	450	24	1	2.0	0	15300	27300
112	11	30	13918823	490	24	1	2.0	0	7800	4500
113	11	30	14398474	374	24	1	2.0	0	400	17800
114	11	30	12758486	351	24	1	1.5	0	1750	800
116	2	18	5885340	420	21	1	2.0	0	1200	3400
117	2	28	5189223	270	24	1	2.0	0	60	1800

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
108	3	0	0	0	0	0	0	0	0	0	0	0	0	4	6350
109	3	210	27	0	0	0	0	0	0	0	0	0	237	4	6350
110	3	80	83	0	0	0	0	0	0	0	0	0	163	4	6350
111	1	0	95	30	97	138	110	120	35	0	0	0	625	0	6350
112	3	53	16	15	66	0	0	0	0	0	0	0	150	4	6350
113	3	28	57	35	0	0	0	0	0	0	0	0	120	4	6350
114	3	0	0	0	0	0	0	0	0	0	0	0	0	4	6350
116	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2603
117	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
118	2	18	5686894	430	21	1	2.0	0	1100	2500
119	2	18	4813824	460	21	1	2.0	0	3900	3800
120	3	12	4093342	366	21	1	2.0	0	400	600
121	2	28	7821456	438	24	1	2.0	0	1000	13900
122	5	28	5688770	221	24	1	2.0	0	0	1060
123	2	28	6181035	237	24	1	2.0	0	1600	3500
124	5	19	5098395	210	21	1	2.0	0	150	400
126	2	18	5431683	344	21	1	1.5	0	30	4370
127	2	18	6095409	340	21	1	2.0	0	2250	550

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
118	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2603
119	1	91	0	0	0	0	0	0	0	0	0	0	91	0	2553
120	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2221
121	2	58	302	0	0	0	0	0	0	0	0	0	360	2	3720
122	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3453
123	1	0	92	40	35	0	0	0	0	0	0	0	167	0	3670
124	1	41	15	0	0	0	0	0	0	0	0	0	56	0	2536
126	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2603
127	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2603

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
128	5	28	7344552	206	24	1	2.0	0	3950	500
129	2	28	5538262	270	24	1	2.0	0	4800	2050
130	5	19	4938182	250	22	1	1.0	0	700	200
131	2	28	6130537	330	24	1	1.0	0	1100	1050
132	10	21	6931120	375	18	1	2.0	0	1450	580
133	5	28	8541311	233	24	1	2.0	0	650	200
134	2	18	4673077	100	21	1	1.0	0	25	5000
135	2	18	5437666	330	21	1	2.0	0	400	7300
137	2	18	6083932	340	21	1	2.0	0	130	6700

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
128	1	23	34	31	50	24	17	0	0	0	0	0	179	0	3453
129	1	69	17	24	8	17	44	19	0	0	0	0	198	0	3670
130	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2536
131	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
132	2	0	0	0	0	0	0	0	0	0	0	0	0	2	3870
133	1	15	0	5	0	20	20	10	20	10	10	55	165	0	3453
134	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2553
135	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2553
137	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2553

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
140	2	18	6045409	290	21	1	2.0	0	4600	1250
141	2	18	6529180	3664	21	1	2.0	0	100	2450
142	3	12	4942961	340	18	1	2.0	0	4100	2300
143	3	20	6960810	276	24	1	2.0	0	6550	470
144	3	20	6390231	233	24	1	2.0	0	1000	2450
145	2	28	7337905	340	24	1	1.5	0	200	3700
146	2	26	8335448	330	24	1	2.0	0	3900	3600
147	2	28	6908921	219	24	1	1.5	0	950	550
148	5	19	5275020	170	21	1	2.0	0	150	1050

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
140	2	0	31	25	78	0	0	0	0	0	0	0	134	2	2603
141	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2603
142	1	35	0	0	0	0	0	0	0	0	0	0	35	0	2221
143	3	40	40	35	20	15	0	0	0	0	0	0	150	4	3220
144	3	85	35	0	0	0	0	0	0	0	0	0	120	4	3220
145	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
146	1	50	105	26	17	0	24	0	0	0	0	0	222	0	3670
147	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
148	1	0	41	0	0	0	0	0	0	0	0	0	41	0	2536

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
149	2	28	6957923	255	24	1	1.0	0	8200	0
150	2	28	10990783	565	24	2	0.6	1	1500	10700
151	2	28	7577367	295	24	1	2.0	0	5600	13100
152	2	18	8850205	310	21	2	0.5	1	50	2900
153	10	21	7143186	350	18	1	2.0	0	7400	750
154	3	12	7408094	360	24	1	2.0	0	1900	1400
155	2	18	5118840	190	21	1	1.0	1	1800	4200
156	11	20	9995390	325	24	1	2.0	0	350	6000
157	2	18	4350750	270	21	1	2.0	0	0	2350

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
149	1	0	72	30	56	26	0	0	0	0	0	0	184	0	3670
150	1	140	0	0	0	0	0	0	0	0	0	0	140	0	3670
151	1	0	38	0	78	0	69	0	44	0	26	28	283	0	3670
152	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2553
153	2	74	38	27	14	6	2	3	0	0	0	0	164	2	3870
154	3	0	120	0	0	0	0	0	0	0	0	0	120	4	3220
155	1	60	55	0	0	0	0	0	0	0	0	0	115	0	2553
156	1	80	80	0	0	0	0	0	0	0	0	0	160	0	5600
157	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2553

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
158	11	25	11810756	350	24	1	2.0	0	4100	10600
159	11	20	10643880	425	24	1	2.0	0	600	5500
160	11	25	11684579	340	24	1	2.0	0	14500	950
161	2	18	5932772	320	21	1	2.0	0	3150	710
162	3	12	7078125	331	24	1	1.0	0	3100	1770
163	2	28	6156745	270	24	1	2.0	0	5200	2000
164	2	28	6749661	318	24	1	2.0	1	6000	1400
165	2	18	5203988	240	21	1	1.5	0	120	4450
166	2	28	6873643	348	24	1	2.0	1	2050	8600

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
158	2	105	30	50	15	20	0	0	0	0	0	0	220	2	6350
159	1	0	0	0	0	0	0	0	0	0	0	0	0	0	5600
160	1	0	43	24	10	15	78	0	0	0	0	0	170	0	6350
161	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2603
162	3	20	0	0	0	0	0	0	0	0	0	0	20	4	3220
163	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
164	1	37	0	0	0	0	0	0	0	0	0	0	37	0	3670
165	2	0	0	0	0	0	0	0	0	0	0	0	0	2	2603
166	1	113	5	0	0	0	0	0	0	0	0	0	118	0	3670

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
167	3	22	5059463	330	24	1	2.0	0	2600	900
168	3	22	6008365	306	24	1	2.0	0	3400	1950
169	2	18	4587976	306	21	1	1.0	0	200	6700
170	2	28	4761702	235	24	1	2.0	0	1100	500
171	2	28	5251840	315	24	1	1.0	0	400	1550
172	2	28	5342725	225	24	1	2.0	0	100	2400
173	2	28	5103970	282	24	1	1.0	0	500	1450
174	2	28	5695288	304	24	1	2.0	0	3600	1130
175	2	28	5463114	230	24	1	2.0	0	300	950

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
167	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3172
168	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3172
169	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2553
170	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
171	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
172	2	0	0	0	0	0	0	0	0	0	0	0	0	2	3720
173	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
174	1	55	10	23	0	0	0	0	0	0	0	0	88	0	3670
175	2	0	0	0	0	0	0	0	0	0	0	0	0	2	3720

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
176	2	28	5991870	270	24	1	1.5	0	950	900
177	2	28	5953221	270	24	1	1.0	0	5	2000
178	2	28	6935347	302	24	2	0.7	1	750	850
181	2	18	4036257	230	21	1	1.0	0	250	1250
182	5	28	5297435	220	24	1	2.0	0	20	2900
183	2	18	5986049	322	21	1	2.0	0	8850	950
184	2	28	6259610	309	24	1	1.5	0	0	5500
185	3	12	4645076	242	18	1	2.0	0	25	1700
186	2	28	6007055	310	24	1	2.0	0	2500	730

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW.700	TOTAL RW	LAB NO	CONCRETE AREA
176	2	0	0	0	0	0	0	0	0	0	0	0	0	2	3720
177	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
178	2	0	0	0	0	0	0	0	0	0	0	0	0	2	3720
181	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2553
182	1	35	35	0	0	0	0	0	0	0	0	0	70	0	3453
183	2	87	12	10	40	7	4	0	0	0	0	0	160	2	2603
184	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670
185	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2221
186	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3670

RAW DATA TABLE

OBS. NO.	MODEL #	NO. OF CLASS-ROOMS	BID PRICE	LEGTH OF FENCE	EXCUTION TIME	FOOTING TYPE	BEARING CAPACITY	SOIL IMPROVEMENT	EXCAVATION	FILL
187	5	19	4915859	225	24	1	1.0	1	0	1400
189	3	12	4724756	225	18	1	2.0	0	200	1400
190	3	20	7389708	243	24	1	1.0	0	0	2700
192	2	28	7036454	290	24	1	2.0	0	1350	3150
193	2	18	7386606	321	21	1	2.0	0	13350	3250
194	11	30	13722367	470	24	1	2.0	0	4500	43400
195	11	30	13288946	450	24	1	2.0	0	11000	26000
196	11	25	11562856	341	24	1	2.0	0	750	5000

RAW DATA TABLE

OBS. NO	LEVEL	RW 150	RW 200	RW 250	RW 300	RW 350	RW 400	RW 450	RW 500	RW 550	RW 600	RW 700	TOTAL RW	LAB NO	CONCRETE AREA
187	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2536
189	1	0	0	0	0	0	0	0	0	0	0	0	0	0	2221
190	3	0	0	0	0	0	0	0	0	0	0	0	0	4	3220
192	1	0	0	30	50	0	0	0	0	0	0	0	80	0	3670
193	2	35	66	19	0	19	35	130	0	0	50	0	354	2	2603
194	2	66	144	22	256	16	17	0	0	0	0	0	521	2	6350
195	2	0	30	83	55	88	60	0	0	0	0	0	316	2	6350
196	1	56	7	0	0	0	0	0	0	0	0	0	63	0	6350

APPENDIX - D

METHOD OF LEAST SQUARES

There are many methods available for estimating the parameters of the regression model. some are more beneficial than others if one needs to insure that the individual parameters are estimated with as much accuracy as possible. Other estimators might be more appropriate if the study demands good prediction of the response variable but not necessarily highly accurate parameter estimates.

The least squares method of regression best complies with the objective, which in this case is to minimize the difference between the actual and the predicted values. A brief account of the method of least squares is presented so as to provide the reader an overview of the procedure.

For purposes of simplicity, the discussions is restricted to a multiple linear regression model with two independent variables. Mathematically, the model is expressed as:

$$Yi = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \varepsilon_i$$

$$i = 1,2,..........N$$

By assuming:

1. The expected value of the error component (ε_i) is zero [E (ε_i) = 0], and the variance of the error component (ε_i) is $\sigma^2[V((\varepsilon_i) = \sigma^2]$, for i = 1,2,......N.

2. $\beta_0, \beta_1......, \beta_k$ are (k + 1) parameters, and $x_{i,1}\ x_{i,2},......, x_{i,k}$ are known constants.

The regression function is given by $E(Yi) = \beta_0 + \beta_1 xi,1 + \beta_2 xi,2$. the nature of this function is illustrated in Figure 1. Since there are now two independent variables, the dependent variable Y becomes the third axis in three-dimensional space; the regression function is a plane.

The interpretation of the regression coefficients β_0, β_1 and β_2 in this two independent variable model are: β_0 is the Y-intercept, the point on the y-axis where it is intersected by the plane; β_1 gives the change in the value of the dependent variable Y when x_1 is incremented by one unit and x_2 is held constant, β_2 is also similarly interpreted. the fitted regression plane $y = b_0 + b_1x_1 + b_2x_2$ is determined by least squares: b_0,b_1 and b_2 are least squares point estimators of β_0, β_1 and β_2, respectively. The least squares function is:

$$LS = \sum_{i=1}^{n} [y1 - (\beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2})]^2 = \sum_{i=1}^{n} \epsilon_i^2$$

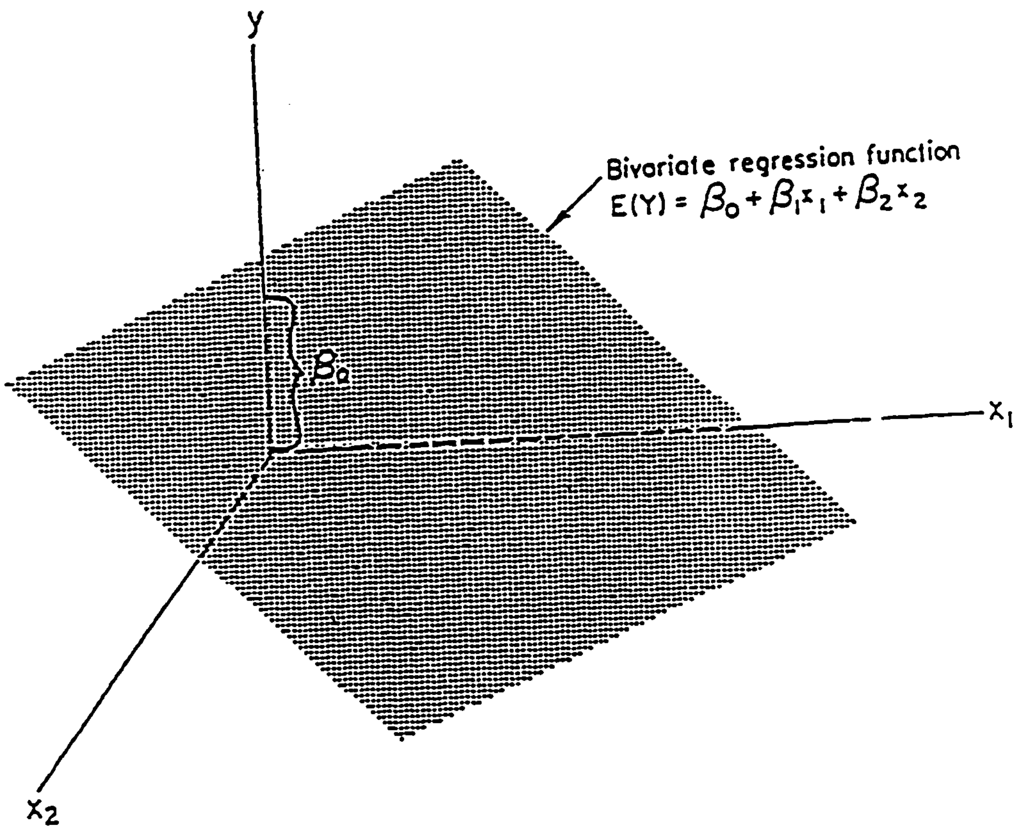

Figure 1: Biyoriate regression function

and the values of β_0, β_1 and β_2, that minimize LS are the expected values of the least squares estimators, b_0,b_1 and b_2 respectively.

The procedure for determining algebraically, the point estimates of the regression coefficients β_0, β_1 and β_2, and the coefficient of multiple determination for a bivariate regression model is presented in the following sections.

POINT ESTIMATION OF THE POPULATION REGRESSION PARAMETERS

Point estimation of the regression coefficients, β_0, β_1 and β_2,: The value of b_0, b_1 and b_2 that minimize LS are solutions to the following normal equations:

$$\Sigma y = nb_0 + b_1 \Sigma x_1 + b_2 \Sigma x_2$$
$$\Sigma x_1 y = b_0 \Sigma x_1 + b_1 \Sigma x_1^2 + \Sigma x_1 x_2$$
$$\Sigma x_2 y = b_0 \Sigma x_2 + b_1 \Sigma x_1 + b_2 \Sigma x_2$$

These equations can be solved algebraically for b_0, b_1 and b_2, but the resulting formulas are quite complicated. Usually, these equations are solved on a computer by using a multiple regression computer program.

POINT ESTIMATION OF THE COEFFICIENT OF MULTIPLE DETERMINATION

The relative strength of the linear relationship between the dependent variable and the set of independent variables is measured by the coefficient of determination, R^2. The coefficient of multiple determination is the proportion of variability in the dependent variable Y accounted or explained by the independent $x_1, x_2, \ldots, x_k$. It is a point estimate of the population coefficient of determination. The value of R^2 is given by the expression:

$$R^2 = \frac{SST - SSE}{SST}$$

where

SST = Sum of squares total

SSE = Sum of squares due to error

The coefficient of determination R^2 can also be expressed as:

$$R^2 = \frac{\text{sum of squares explained by regression}}{\text{total sum of squares}}$$

OR

$$R^2 = \frac{\text{explained variation}}{\text{total variation}}$$

It may be noted that R^2 can take any value between zero and one. Larger values of R^2 indicated a stronger relationship and smaller values a weaker one.

TEST OF THE SIGNIFICANCE OF THE MODEL - (THE F TEST)

To test in the significance of the model developed, a joint test of the regression coefficients is performed. This test consists of testing the following hypothesis.

Null Hypothesis:

H:0 $\beta_0 = \beta_1 = \beta_2 = \beta_3 = \ldots\ldots\ldots \beta_k = 0$

The null hypothesis H0 states that all regression parameters equal zero which implies that no relationship exists between the variables in question and that the set of regression parameters to estimate the value of Y. Depending on the statistic test there are two outcomes. Either it is said that the null hypothesis 'H0' is REJECTED or it is said that the null hypothesis 'H0' CANNOT BE REJECTED. The test concerning the significance of the regression relationship is based on the following F statistic:

$$\text{F-ratio} = \frac{MSR}{MSE}$$

Where MSR is the 'mean square due to regression' and MSE is the 'mean square due to error.' The F-ratio is then compared with the critical F-value which is F $(1-\alpha;\ k,n-k-1)$, where 'α' is the specified level of significance (.01-.1), 'k' is the number of independent variable and 'n' is the number of observations. this value is read from F-distribution tables which can be found in any standard statistics text. If the F-ratio is greater than F $(1-\alpha;k,n-k-1)$, then it is said that enough evidence is available to reject the null hypothesis H0 at the 'α' level of significance. Consequently, if the F-ration is less than, F $(1-\alpha;k,n-k-1)$, then it is inferred that there is not sufficient evidence to reject the null hypothesis H_0 at the 'α' level of significance.

RESIDUAL

A residual is the difference between an observed Y and the corresponding Y predicted by the fitted linear relationship. For a linear regression model the residuals (which are the same as errors), are assumed to be normally distributed with a mean of zero and a constant standard deviation. Furthermore, they are assumed to be statistically independent. Residuals can be used to test and verify these assumptions. A bell-shaped distribution of the frequency histogram for the residuals would be supportive of the normality assumption. However, formal tests for normality are not often performed. In large samples, lack of normality has no important consequence, and in small samples it is difficult to prove.

REFERENCES

1. Al-Asfoor, M.D. (1993). "Computerized Conceptual-Cost-Estimating System for Domestic Water Supply Projects", M.S. Thesis, King Fahd University of Petroleum & Minerals, Dhahran, Saudi Arabia.

2. Albright, Christian S. (1987). Statistics for Business and Economics, Macmillan, New York.

3. Annual Statistics Book (27th Edition). (1991). K.S.A. Ministry of Finance.

4. Bozai, G.A. (1981). "Regression Analysis Models for Preliminary Cost Estimating in High Way Construction", MS Thesis, NED University of Engineering and Technology, Pakistan.

5. Building Construction and Data. (1981). 39th ed. Kingston: Robert Snow Means Co., pp.284-285.

6. Dagosiino, F. (1989). Estimating in Building Construction, Prentice Hall, Englewood Cliffs, N.J.

7. Dodge Construction Systems Costs. (1981). New York: MacGraw-Hill, pp. 7-16.

8. Dodge Construction Cost Information System. (1984). New York: MacGraw-Hill, Volume 2.

9. Fifth Development Plan. (1990 - 1995). K.S.A. Ministry of Planning.

10. Helton, Joseph E. (1985). Simplified Estimating for Builders and Engineers, Prentice-Hall, London.

11. Keith, C (1987). Fundamentals of Construction Estimating and Cost Accounting, Prentice Hall, Englewood Cliffs, N.J.

12. Neter, J., and Wasserman, W., and Whitemore, G.A. (1978). Applied Statistics, Allynward Bacon, Boston.

13. Pubier, H.E. (1969). Construction Estimates and Costs, MacGraw Hill, New York.

14. Raddon, P.S.B, (1982). Building Cost Techniques, E. and F.N., Spon, London.